

## FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT  
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA  
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC  
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC  
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG  
AGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA  
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT  
TCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT  
GTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCC  
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTATAGGACCGCCTAC  
CGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG  
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG  
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG  
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC  
CGGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG  
TGCCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG  
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT  
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC  
TCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTTC  
CTGGAGGAGCAGCTGGGGTCTTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGG  
CTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTC  
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCTCCTC  
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT  
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA  
GCTGAGGGAAGGTACGAGTTCCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC  
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAAATGAAA  
CGTGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCT  
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT  
TACAAAT

## **FIGURE 2**

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV  
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC  
GAAICQPPCRNGGSCVQPGRRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCW  
EGHSLSADGTLCVPGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLAS  
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

### **Signal sequence:**

amino acids 1-19

### **cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 93-97, 270-274

### **N-myristoylation sites.**

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,  
191-197, 265-271

### **Amidation site.**

amino acids 26-30

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 152-164

### **Cell attachment sequence.**

amino acids 130-133

### **EGF-like domain cysteine pattern signature.**

amino acids 123-135

### FIGURE 3

CGCTCGCCCCGTCGCCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG  
GTCAGCCCACGGCGGGGACT**ATG**GTGAAATTCCCGGCGCTCACGCACTACTGGCCCCCTGATC  
CGGTTCTTGGTGCCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA  
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT  
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGCCTGGTGTTT  
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCCTGTGTATGGTGGTGGCAGGGGCCATCGC  
TGCCGTCCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC  
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT  
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAAACACAAATACAGTTT  
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGTAGCCATTTTGC  
TTCACAGTCACCTGGAATGCCGGGAGCCCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC  
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTACGACATCATCCC  
TGACAGAAGTGGCCCCGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT  
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT  
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA  
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG  
ACAAGAATAACCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCCACATC  
AAGAAGTTACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC  
ACCCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC  
TCTGTGTTGTTTCCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT  
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGCCCCCAGCTCTGTGCTGCG  
GATCATCGTCCTCATCGCCAGCCTCGTGGTCTACCTACCTGGGGGTGCACGGTGCAGCCC  
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTTCGCCATCGCTGCG  
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA  
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAATGAGAGAGG  
AGAATGAAT**TAA**GGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC  
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTTGTTTTGTTTTGGTAAT  
GAAAGAGGCCTTGATTTAAAGGTTTTCTGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT  
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC  
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC  
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT  
CACCTGACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG  
TAAAACTCGGCTTCCTTTGATTTGCTTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC  
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG  
GAGGCGGTGGCACGCTGCAGCCCGGAGTCCCCGTTCACTGAGGAACGGAGACCTGTGAC  
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG  
GGCAGCAAACTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA  
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTACAAAA  
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCTGAAGGCTTGCTTTTCCCTCGC  
CTTTCCTGAAGGTCGCATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT  
TACAGTGAAGCTTAAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC  
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA  
TTGAGAATGTACTACGGTACTTCCCTCCACACCATAACGATAAAGCAAGACATTTTATAACG  
ATACCAGAGTCACTATGTGGTCCTCCCTGAAATAACGCATTTCGAAATCCATGCAGTGCAGTA  
TATTTTCTAAGTTTTTGAAAGCAGGTTTTTTTCTTTAAAAAATTATAGACACGGTTCACT  
AAATTGATTTAGTCAGAATTCCTAGACTGAAAGAACCTAAACAAAAAATATTTTAAAGATA  
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTTCTATTTTCGC  
ATTTTCAATAAAATGTCTCTAATAACAAAAA

1000 950 900 850 800 750 700 650 600 550 500 450 400 350 300 250 200 150 100 50 0

## **FIGURE 4**

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF  
TGPMSDFKNVGLVVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDESV  
GSKTRRAFLYLAAFPMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC  
REPLLIPILSLYMGALVRCTTLC LGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL  
ATQORISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN  
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelcVVPLR  
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVHGATLGVGSL  
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

### **Transmembrane domains:**

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,  
408-423, 431-445

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.





### FIGURE 7

TATTTCCAGTTCCGGTTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA  
ACCTTNGTCCTTGCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT  
CCTACCCTACCTGGGGGTGCACGGTGAGAC

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

## FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA  
GCCTGCTCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG  
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCC  
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG  
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG  
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT  
TCTTTTTTACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG  
AATGGGTTTTTGGTTCTTTAAGTTCTCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT  
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCA  
TCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC  
AAGGCCGAGGAGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT  
CTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCCAGCGGCT  
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT  
GCTGTCTTGCCCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTTCAT  
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA  
ACCCCCATTTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG  
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT  
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC  
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCC  
TTTGACAACGAGCAGGACGGCGTCACTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT  
GGCCTCACTGCACGTTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA  
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC  
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCT  
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT  
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTG  
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA  
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCCTCCTCCTGT  
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCCTCAGGCTCCACGGGAGCGGGG  
CTGCTGGAGAGAGCGGGGAACCTCCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT  
CCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTTCTGGAATTTCGTGCCTTACTGAGTCTCT  
AAGACTTTTTTCTAATAACAAGCCAGTGCGTGTAACAAAAA

1000 900 800 700 600 500 400 300 200 100

## **FIGURE 9**

MGACLGACSLLSASCCLCGSAPCILCSCCPASRNSTVSRIFTFFLFLGVLVSIIMLSPGVE  
SOLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFFTLLMLCVSSSRD  
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSW  
NQRWLGKAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC  
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA  
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA  
ACEGRAFDNEQDGVITYSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICAS  
WAGLLLYLWTLVAPLLLRNRDFS

### **Signal sequence:**

amino acids 1-20

### **Transmembrane domains:**

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGAGCCCTCTGGCAGAGGGTTAACCTGGGTC  
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA  
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTGAGAGAAGGCCATGGAGGTGCCGCCACCGGC  
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG  
TGA CTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC  
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG  
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG  
TGTATGGGGGAATACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA  
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT  
CATTTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG  
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAATTGCAGGA  
GCTGTACGGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT  
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG  
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAAC TGAA  
GAGTGGAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG  
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC  
CTTCAGTAATAGATAAACAAGACAAGGACTTGAAAGTGCTCTGAACTTGAAACTCACTGGAGA  
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC  
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTTCTTTCTT  
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT  
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA  
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

100529533460

## **FIGURE 11**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAIEYHNRFDAVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGS�FRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD
```

### **Important Features:**

#### **Signal Peptide:**

amino acids 1-24

#### **Transmembrane domains:**

amino acids 76-96 and 171-195

#### **N-glycosylation site:**

amino acids 153-156

## **FIGURE 12**

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA  
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA  
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA  
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT  
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTCGTTTCATG  
GCTGGCGCCGAACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

1. The first group of people who are interested in the study of the history of the United States are the students of the history of the United States.

## FIGURE 14

GAGCCGCCGCCGCGCGCGCGCGCGCGCACTGCAGCCCCAGGCCCCGGCCCCCACCACGTCT  
GCGTTGCTGCCCCGCCTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT  
CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGTTTGCAGGGG  
GCTTCGCGTGTTCCAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGTTAGTCTG  
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT  
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCTGATTGCTTTAGTGGGTCTGATTGGAG  
CTGTAAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTATT  
GTTTCAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT  
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAACTGCT  
GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAGTGACCAC  
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG  
TGGCATTGGCCTGTTCTTCAGTTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA  
GGAACCAGAAAGACCCCCGCGCGAATCCTAGTGCATTCTTTTGAAGAGAAAACAAGGAAGAT  
TTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTCTGTAAAGCTCCATTTGCCAGT  
TTAAGGAAGGAAACACTATCTGGAAGGTAACCTTATTGATAGTGAATTATATATTTTTACT  
CTATGTTTCTCTACATGTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAACTTGTGGTCTC  
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTGCGGCACTGTCCACTGTGGCCTT  
TCTTAGCATTTTTTACCTGCAGAAAACTTTGTATGGTACCCTGTGTTGGTTATATGGTGAA  
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC  
TGGAAGGAGAGTGGAAATTTATTAATAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA  
TCCAAATCCCAATTTTTTTTTGGTCTTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA  
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT  
GTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT  
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT  
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA  
AAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGATATTTG  
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT  
GTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATTGTGTG  
GCTCTGTATATTCTGTAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA  
ATTAAAGAAAGTAATGGAAG

## **FIGURE 15**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALV  
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQR  
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRVGGIGLFFSFTEILGVWL  
TYRYRNQKDPRANPSAFL

### **Signal Peptide:**

amino acids 1-34

### **Transmembrane domains:**

amino acids 47-63, 72-95 and 162-182

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151

## FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT  
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG  
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT  
NTAAACTGCTGTGGGTTCGGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA  
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAAATATGCTGGAGAGGTTTTGA  
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC  
TACAGATACAGGAACCAG

60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

## FIGURE 18

ATGATTATTCTGTACTTGTATTTATTGTTTCTGCGCTTGTTTAGCCC  
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG  
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC  
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA  
TCATAGGAGAATATGC

100-20-50-100

## FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG  
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTGTGAGACGCTGCAGTGTGAGGGACCTGT  
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT  
TCCAGGTCAAGGCCTACACTTTCAGTGAACCCTTCCACCTGATTGTGTCCTATGACTGGCTG  
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC  
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC  
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC  
TGCAGTGGCATCTTCCAGAGCCCTGGTCCTGGGATCCAGAAACAGCATCTGTTGTGGCTAT  
CACAGTCCAAGAACTGTTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG  
CAGGAAGCCCCATGACCCTGAGTTGTGAGACAAAGTTGCCCTGCAGAGGTGAGCTGCCCCG  
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT  
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCATACTGGTGTGAGGCAGCCACTGAGG  
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT  
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA  
GGAGGCCCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT  
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCGATGGGCCTTCTTCTCAAACACATG  
CAGGATGTGAGAGTCCTCCTCGGTCACTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA  
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAAGTAAACAGTTCATCCATGATCTCACT  
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTGTCTGACATATGCATAAGTA  
CTTTTACAAGTTGTCCAGTGTTTTGTGTAGAATAATGTAGTTAGGTGAGTGTAATAAATTT  
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC  
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA  
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA  
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTGGGGGATGGTGGGATGT  
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG  
GGCAGTACCCCAACAACGAAAAATAATCTGGCCCAAAATGTCAGTTGTACTGAGTTTGAGAAA  
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT  
TATCTCTTTCCAGCCTCATTGAGCTATTCTTACTGACATAACAGTCTTTAGCTGGTGCTATG  
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAAGCCATTATGTTGAAATCCTAATCCCC  
AAGGTGATGGCATTAAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC  
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCTTCCACCAT  
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCACAAAAACAGCTGTCGCCAACACCG  
ACTCTGTCGTTGCCTTGATCTTGAACCTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG  
TTGTTTGTAGCCTAA

## **FIGURE 20**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV  
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRCAWQDWPLTQVTFYRDGSALGPPGP  
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS  
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ  
VWKQSPQLEIRVQGAASSAAPTLPNPAPQKSAAPGTAPPEAPGPLPPPPTPSSDPGFSSPL  
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

### **Signal sequence:**

amino acids 1-17

### **Leucine zipper pattern sequence:**

amino acids 12-33

### **Protein kinase C phosphorylation site:**

amino acids 353-355

## FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT  
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTG**ATG**GGG  
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT  
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG  
ACCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT  
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG  
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCCTGGAGATGG  
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG  
AGAGATAAGATTACTGAGCTCCGTGTCCAGAACTCTCTGTCTCCAAGCCCACAGTGACAAC  
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC  
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC  
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA  
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG  
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC  
TTGAAAGCAACATCTACAGTGAAGCAGTCTGGGACTGGACCACTGACATGGATGGCTACCT  
TGGAGAGACCAGTGTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT  
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTTCGGAAGACATCCCAA  
CAAGAGCATGTCTACGAAGCAGCCAGG**TA**AGAAAGTCTCTCCTCTTCCATTTTTTGACCCCGT  
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGTGTGGCACAGACCC  
AATCCTAAGGCCGGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT  
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC  
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA  
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA  
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT  
CCAGTGATGAGCCAACTTCCCAGAATCTGGGCAACAACACTCTGATGAGCCCTGCATAGGA  
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC  
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAGTGTCTGTTAAAAATGCCCCATTAGGC  
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT  
ACCTCTCTTCTGGATAGCCCCAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT  
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG  
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG  
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA  
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT  
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTACAGGCCAGGGTTCA  
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA  
ATCATAACAGC

7002055400

## **FIGURE 22**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLGHLLTVDTYGRPILEVPESVTGPWKGDVNLPTDYDPLQGYTQVLVKWLVQRGS  
DPVTIFLRDSSGDHIQQAKYQGRLLHVSHKVPGDVSLQLSTLEMDDRSYHTCEVTWQTPDGNQ  
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRLSLQCQARGSPPISYIWKQQTNNQE  
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFFVKDSSKLLKTKTEAPTTMT  
YPLKATSTVKQSWDWTDDMDGYLGETSAGPGKSLPVFAIILIIISLCCMVVFTMAYIMLCRKT  
SQQEHVVEAAR

### **Signal Sequence:**

amino acids 1-19

### **Glycosaminoglycan attachment site:**

amino acids 149-152

### **Transmembrane domain:**

amino acids 282-300

## FIGURE 23

GCGCCGGGAGCCCATCTGCCCCCAGGGGCACGGGGCGCGGGGCGCGGCTCCCGCCCCGGCACAT  
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCGCGCCAGCTCGCCCGAGGTCCGTGCGA  
GGCGCCCGGGCCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC  
GGGATGTCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA  
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC  
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA  
GTGGTGATCACTTACTCCAGTCGTATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG  
AGTGGCCTTTGCTTCCAATTTCTTGGCAGGAGATGCCTCCTTGACAGATTGAACCTCTGAAGC  
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT  
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC  
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT  
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT  
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA  
CCAGTGACACAGCAGGCAACGAAGCTGGGAAGGAAAAGCTGTGTGGTGCGAGTAACTGTACAGT  
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG  
ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA  
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT  
CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT  
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC  
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAAGTCCACCATG  
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA  
ACGGTCTTGAATTACAATGGACTTGACTCCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC  
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA  
GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTCCTTATACAATACCAAACAAGCAAA  
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG  
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG  
AGGTGAATATACCTAAAACCTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT  
TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTCTATGCATTTCTGCAAACCTATTGGATT  
ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC  
TGAGCTAACCACCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC  
TTCATTTGTCTATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA  
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAATAAC  
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTCCATCTTCATGATGTT  
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCTCAAAT  
CAGATGCCTCTAAGGACTTTCTGCTAGATATTTCTGGAAGGAGAGAAAATACAACATGTCATT  
TATCAACGTCCTTAGAAAGAATCTTCTAGAGAAAAGGGATCTAGGAATGCTGAAAGATTA  
CCCAACATACCATTATAGTCTCTTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG  
GTGGACTAGAAAGGGAGATTAGATCAGTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA  
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC  
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

**THE** **NEW** **YORK** **PUBLIC** **LIBRARY**

><subunit 1 of 1, 373 aa, 1 stop

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV  
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV  
ILKVLVRPSKPKCELEGELTEGSDLTLCCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID  
YNHPGRVLLQNLTMSSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI  
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSSGSRSSRSRSSSTRSTANS  
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

amino acids 1-16

## amino acids 232-251

## FIGURE 25

GTCGTTCTTTGCTCTCTCGCGCCAGTCTCTCCTCCCTGGTTCTCCTCAGCCGCTGTCGGAGGAGAGCACCCGGA  
 GACGCGGGCTGCAGTCGCGGCGGCTTCTCCCGCCTGGGCGGCCTCGCCGCTGGGCAGGTGCTGAGCGCCCCTAG  
 AGCCTCCCTTGCCGCTCCCTCCTCTGCCCCGCGCAGCAGTGCACATGGGGTGTTGGAGGTAGATGGGCTCCCG  
 GCCCCGGAGGCGGCGGTGGATGCGGCGCTGGGCAGAAGCAGCCGCGGATTCCAGCTGCCCCGCGCGCCCCGGGCG  
 CCCCTGCGAGTCCCCGGTTTCAGCCATGGGGACCTCTCCGAGCAGCAGCACCGCCCTCGCCTCCTGAGCCGCATC  
 GCCCGCCGAGCCACAGCCACGATGATCGCGGCTCCCTTCTCCTGCTTGGATTCTTAGCACCACCACAGCTCAG  
 CCAGAACAGAAGGCCTCGAATCTCATTGGCACATACCGCCATGTTGACCGTGCCACCGGCCAGGTGCTAACCTGT  
 GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCAGCAGTTGCCCT  
 GTGGGACCTTTACCAGGCATGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCATGGCCAATG  
 ATTGAGAAATTACCTTGTGCTGCCCTGACTGACCGAGAATGCACTTGCCACCTGGCATGTTCCAGTCTAACGCT  
 ACCTGTGCCCCCATACGGTGTGTCTGTGGGTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG  
 TGTAAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCCTTCTAGTGTGATGAAATGCAAAGCATACACAGACTGT  
 CTGAGTCAGAACCTGGTGGTGTCAAGCCGGGACCAAGGAGACAGACAACGCTCTGTGGCACACTCCCGTCTTTC  
 TCCAGTCCACCTCACCTTCCCCTGGCACAGCCATCTTCCACGCCCTGAGCACATGGAAACCCATGAAGTCCCT  
 TCCCTCACTTATGTTCCCAAAGGCATGAACCTCAACAGAATCCAACCTCTTCTGCCTCTGTTAGACCAAAGGTACTG  
 AGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCCTC  
 CCAACCTTCAGGTAGTCAACCACCAGCAAGGCCCCACCACAGACACATCCTGAAGTGTGCGGTCCATGGAG  
 GCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG  
 CATTTTGACATCAATGAGCATTGCCCCGATGATTGTGCTTTTCTGCTGCTGGTGCTTGTGGTGATTGTGGTG  
 TGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCGGCGAGGATCCAGTGCCATTGTGGAAAAGGCA  
 GGGCTGAAGAAATCCATGACTCCAACCCAGAACCAGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT  
 ATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAG  
 AGGGAGGTGTGCTGCTTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGACC  
 ATCCGGGGCCCCGAGGCCAGCCTCGCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAG  
 AAGATTTCGTGGGCTGATGGAAGACACCACCAGCTGGAAACTGACAAACTAGCTCTCCGATGAGCCCCAGCCCG  
 CTAGCCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCGCTCTCCTGACGGTGGAGCCTTCCCCA  
 CAGGACAAGAACAAGGGCTTCTCGTGGATGAGTCGGAGCCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCC  
 TCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACACAGTGTGCGGCAGGTACGCCTGGAC  
 CCCTGTGACTTGCAGCCTATCTTTGATGACATGTCCACTTCTTAAATCCTGAGGAGCTGCGGTGATTGAAGAG  
 ATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTGCAAAATTATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC  
 CTCTGGACTCTGTTTATAGCCATCTTCTGACCTGCTGTAGAACATAGGGATACTGCATTCTGGAAATTACTCA  
 ATTTAGTGGCAGGGTGGTTTTTTTAAATTTCTCTGTTTCTGATTTTTGTTGTTTGGGGTGTGTGTGTGTGTTGT  
 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTAAACAGAGAATATGGCCAGTGTGAGTTCTTCTCCTTCTC  
 TCTCTCTCTTTTAAAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAAGTGTGTGAA  
 ATACCCACCACTAAAGTTTTTTAAGTTCATATTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATTCTG  
 TGCATTTAAATTTACTTAACTTACCATAAATGCAGTGTGACTTTTCCACACACTGGATTGTGAGGCTCTTAAC  
 TTCTTAAAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTTATGTCTCTTAACATTACACCTACTTTTT  
 AAAAACAAATATTATTACTATTTTTATTATTGTTTGTCTTTATAAATTTCTTAAAGATTAAGAAAATTTAAGA  
 CCCATTGAGTTACTGTAATGCAATTCACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTTTATATTTCATGG  
 CTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTCACCTGGACACCGTGTAGAATGCTTGATTACTTGTAC  
 TCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT  
 GACAACTGGGCCACCAAGAAGCTTGAACCTTACCTTTTAGGATTGAGCTGTTCTGGAACACATTGCTGCACCTT  
 GGAAAGTCAAAATCAAGTGCCAGTGGCGCCCTTTCATAGAGAATTTGCCAGCTTTGCTTTAAAAAGATGTCTTG  
 TTTTTTATATACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCTGGTGGGATTCTTTCACCAATT  
 ACTTTAATTAAAAATGGCTGCACTGTAAGAACCCTTGTCTGATATATTTGCAACTATGCTCCCATTTACAAATG  
 TACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCAAAGGTGGCGTGGACTCCCTTGTGTGGGTGGGGTTTGTGG  
 GTAGTGGTGAAGGACCGATATCAGAAAAATGCCTTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTTA  
 AAAAAAAA

1000 900 800 700 600 500 400 300 200 100 0

Country	Year	Value	Unit
Algeria	1990	1.00	1000
Algeria	1991	1.00	1000
Algeria	1992	1.00	1000
Algeria	1993	1.00	1000
Algeria	1994	1.00	1000
Algeria	1995	1.00	1000
Algeria	1996	1.00	1000
Algeria	1997	1.00	1000
Algeria	1998	1.00	1000
Algeria	1999	1.00	1000
Algeria	2000	1.00	1000
Algeria	2001	1.00	1000
Algeria	2002	1.00	1000
Algeria	2003	1.00	1000
Algeria	2004	1.00	1000
Algeria	2005	1.00	1000
Algeria	2006	1.00	1000
Algeria	2007	1.00	1000
Algeria	2008	1.00	1000
Algeria	2009	1.00	1000
Algeria	2010	1.00	1000
Algeria	2011	1.00	1000
Algeria	2012	1.00	1000
Algeria	2013	1.00	1000
Algeria	2014	1.00	1000
Algeria	2015	1.00	1000
Algeria	2016	1.00	1000
Algeria	2017	1.00	1000
Algeria	2018	1.00	1000
Algeria	2019	1.00	1000
Algeria	2020	1.00	1000
Algeria	2021	1.00	1000
Algeria	2022	1.00	1000
Algeria	2023	1.00	1000
Algeria	2024	1.00	1000
Algeria	2025	1.00	1000
Algeria	2026	1.00	1000
Algeria	2027	1.00	1000
Algeria	2028	1.00	1000
Algeria	2029	1.00	1000
Algeria	2030	1.00	1000
Algeria	2031	1.00	1000
Algeria	2032	1.00	1000
Algeria	2033	1.00	1000
Algeria	2034	1.00	1000
Algeria	2035	1.00	1000
Algeria	2036	1.00	1000
Algeria	2037	1.00	1000
Algeria	2038	1.00	1000
Algeria	2039	1.00	1000
Algeria	2040	1.00	1000
Algeria	2041	1.00	1000
Algeria	2042	1.00	1000
Algeria	2043	1.00	1000
Algeria	2044	1.00	1000
Algeria	2045	1.00	1000
Algeria	2046	1.00	1000
Algeria	2047	1.00	1000
Algeria	2048	1.00	1000
Algeria	2049	1.00	1000
Algeria	2050	1.00	1000
Algeria	2051	1.00	1000
Algeria	2052	1.00	1000
Algeria	2053	1.00	1000
Algeria	2054	1.00	1000
Algeria	2055	1.00	1000
Algeria	2056	1.00	1000
Algeria	2057	1.00	1000
Algeria	2058	1.00	1000
Algeria	2059	1.00	1000
Algeria	2060	1.00	1000
Algeria	2061	1.00	1000
Algeria	2062	1.00	1000
Algeria	2063	1.00	1000
Algeria	2064	1.00	1000
Algeria	2065	1.00	1000
Algeria	2066	1.00	1000
Algeria	2067	1.00	1000
Algeria	2068	1.00	1000
Algeria	2069	1.00	1000
Algeria	2070	1.00	1000
Algeria	2071	1.00	1000
Algeria	2072	1.00	1000
Algeria	2073	1.00	1000
Algeria	2074	1.00	1000
Algeria	2075	1.00	1000
Algeria	2076	1.00	1000
Algeria	2077	1.00	1000

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594
```

```
><subunit 1 of 1, 655 aa, 1 stop
```

><MW: 71845, pI: 8.22, NX(S/T): 8

[illegible]

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

[illegible]

ATGGGAAGCCAGTAACACACTGTGGCCCTACTATCTCTTCCGTGGTGCCATCTACATTTTGGGA  
CTCGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC  
**CATG**GGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCGGATCGCTTTTGGCC  
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG  
TCACTGCTGCCATTGAAGTTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC  
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT  
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC  
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTACAGCTGCTTC  
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCCAAC  
TGGGTTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC  
CGGGAGGAGTTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA  
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCACTGCACAGCCT  
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAACATGTCCTTGCTCTCGCAG  
TGGCCCTGGCAGGCCAGCCTTCAGTTCAGGGCTACCACCTGTGCGGGGGCTCTGTATCAC  
GCCCCGTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA  
CCATCCAGGTGGGTCTAGTTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGGTGGAGAAG  
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT  
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGA  
ACTTCCCCGATGGAAAAGTGTGCTGGACGTGAGGATGGGGGGCCACAGAGGATGGAGGTGAC  
GCCTCCCCCTGTCCTGAACCACGCGGCCCTCCCTTTGATTTCACAACAGATCTGCAACCACAG  
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG  
TGGACAGCTGCCAGGGGGACAGCGGGGGGGCCCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG  
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC  
CCGTGTCACCTCCTTCCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCCTAAAAACCT**TGA**A  
GAGGAAGGGGACAAGTAGCCACCTGAGTTTCTTGAGGTGATGAAGACAGCCCGATCCTCCCCCT  
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCTTGAGGCTCTGAGTTCCGGCACCA  
GTAGCAGGCCCGAAAGAGGCACCCTTCCATCTGATTCCAGCACAAACCTTCAAGCTGCTTTTT  
GTTTTTTGTTTTTTTGGAGGTGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGAAA  
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTC AAGCGATTCTCTTGCTCAGCTTCCCCA  
GTAGCTGGGACCACAGGTGCCCCGCCACCACACCCAACTAATTTTTGTATTTTTAGTAGAGAC  
AGGGTTTTACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT  
CAGCCTCCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC  
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCCTTCCCACTGTTCCAT  
CTGGTTTTCTCTCCAGGTCCTTGCAAAATTCCTGACGAGATAAGCAGTTATGTGACCTCACG  
TGCAAAGCCACCAACGACCCACTCAGAAAAGACGCACCAGCCCAAGTGCAGAACTGCAGTC  
ACTGCACGTTTTTCATCTCTAGGGACACGAACCAACCCACCCTTTCTACTTCCAAGACTTAT  
TTTTACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTTCATGATTTCTT  
TGTAGCATTTGGTGCTTGACGTATTATTGTCTTTGATTCCAAATAATATGTTTCTTCCCT  
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAAATCATCCACTGAAA

## FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA  
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVVRVGGQNAVLQVFTAAS  
WKTMCSDDWKGYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSI DHLLPDDKVTALHH  
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT  
PLWIITAHCVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHISKYKPKRLGNDIALMKL  
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPPLISNKICNHR  
DVGGGIISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSF GIGCAEVNKP GVYT  
RVTSFLDWIHEQMERDLKT

### **Signal Peptide:**

amino acids 1-20

### **Transmembrane domain:**

amino acids 240-284

[illegible]

CCACGCGTCCGTCCTAGTCCCGGGGCCAACTCGGACAGTTTGCTCATTATTATTGCAACGGTCAAGGCTGGCTTGT  
GCCAGAACGGCGCGCGCGCGCACGCACACACACGGGGGAAACTTTTTTAAAAATGAAAGGCTAGAAGA  
GCTCAGCGCGCGCGCGCGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGA  
CGCCCGCGCCCGGCTCGGCGCGCGCGTGGGATGGTGACAGCGCTCGCCGCGGGGCCGAGAGCTGCTGCACTGAAG  
GCCGCGGACGATGGCAGCGCGCGCGCTGCCGTGTCCCCCGCCCGCGCCCTCTGTCTCGCCCTGGCCGGTGTCTCT  
GCTCGCGCCCTGCGAGGCCGAGGGGTGAGCTTATGGAACCAAGAGAGCTGATGAAGTTGTCACTGCTCTGT  
TCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTCCCAAGATCATCCAGAAGTGTGAATATTCGACT  
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAA  
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT  
ACGGGGATATTCTGATTACAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAAATGA  
AAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAACTCTTCCACGCAAGAAGCTGAAAACGCT  
CCGGGATCATGTGGATCATACACAACACCAAACTCGTGTGCAAGAAGATGTGTTTCCACCACCCCTCTCAGAC  
ATGGGCAAGAAGGCATAAAAGAGAGACCCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG  
AGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAGTTAAGCAGCGATTAAATAGAGATTGCTAATCACGTTGACAA  
GTTTTACAGACCACTGAACATTTCGGATCGTGTTGGTAGGCGTGGAAGTGTGAATGACATGGACAAATGCTCTGT  
AAGTCAGGACCCATTACACGACCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTCACTCGCAAATCCCA  
TGACAAATCGCAGCTTGTCACTGGGTTTATTTCCAAAGGACCAACCATCGGCATGGCCCCAATCATGAGCATGTG  
CAGCGCAGACCACTGTGGGGGAATGTGATGGACCAATCAGACAACTCCCTTGGTGCAGCCGTGACCTTGGCACA  
TGAGCTGGGCCACAATTTCTGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA  
AGGAGGTGCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTTACGAGTTGCAGCAGGAAGGACTT  
GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTGTTTAACTGCCGGAAGTCAGGGAGTCTTTTCGGGGGCCA  
GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGGAGGTGTGACTGTGGGGAGCCAGGAATGTATGAATCGCTGT  
CTGCAATGCGCACCTGTACTCCTGAAGCCGAGCGTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT  
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC  
TCACTGCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTGAGGATGTGGACGGCTACTGCTACAATGGCAT  
CTGCCAGACTCACGAGCAGCAGTGTGTACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA  
GAGAGTCAATTTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCTTTGCCAAATGCGAGAT  
GAGAGATGCTAAATGTGGAAAAATCCAGTGTCAAGGAGGTGCCAGCTCGGCGCCAGTCATTGGTACCAATGCCGTTTC  
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCGGATTTCTGTGCCGGGGGACCCAGTGTACTTTGGGCGATGA  
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCA  
AAATATTAGTGTCTTTGGGGTTTACAGAGTGTGCAATGCAGTGCACGGCAGAGGGGTGTGCAACAACAGGAAGAA  
CTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACCGGGCCC  
CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTTCAGAGTCCCAAGGAGGCGCGGCCAGGGCCAGGACCCGTGGG  
ATCGAGGAGCATGCGTCTACTGCCTACTGACACTCATCTGAGCCCTCCATGACATGGAGACCGTGACCAGTG  
CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT  
TCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCACGTCTGCAGAAA  
CAGTGCAGGAAGGGCAGCGACTTCTGTGTTGAGCTTCTGCTAAACATGGACATGCTTCAGTGTCTCCTGAG  
AGAGTAGCAGGTTACCATTCTGCGAGGCCCAAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAGTCTGGCCCTTTC  
ACTGAGCCTCCAGCAGCTGGGGGAAAGCAAGGTTGGGCCAGTGTCCTTTCCCACTGACACCTCAGCCT  
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATGTGGCTTTTAGCATTTATTATATGAAAT  
AGCAGGGTTTTAGTTTTTAATTTATCAGAGACCTGCCACCCATTCCATCTCCATCCAAGCAAACCTGAATGGCAA  
TGAAACAAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACCTCTGGCTCTTTGCTGTGGACATGCGTGACCAGC  
AGTACTCAGGTTTGGAGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCAACAACCTTTCATTTAAACAAGTAAGAA  
GTTTAAAAAGTGAAACAATGTAAGAGCCTAACTCCATCCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT  
GAAAT

[illegible]

```
><subunit 1 of 1, 735 aa, 1 stop
```

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGLWIPVKSFDSK  
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCIYHGHVRG  
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA  
AKNVFPFPPSQTWARRHKRETLKATKYVELVIVADNREFQORQGDLEKVKQRLIEIANHVDKF  
YRPLNIRIVLVGVEVWDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG  
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK  
GGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE  
CDCGEPEECMNRCNATCTCTLKPDCAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS  
PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPAGKAPAGICFERVNSAGDPYGN  
CGKVS KSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLOQGGRI LCRGTHVYLG  
DDMPDPGLVLGATKCADGKICLN RQCQNI SVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP  
FCDKFGFGGSTDSGPIRQAEARQEAESNRERGQGOEPVGSQEHASTASLTLI

amino acids 1-28

## 1

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAACGAAA  
ACAAAACAGTTTTTGGGGGTT CAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA  
GTGCAAGGAAGGGCAGGANNTTCTTGTTGAGNTTTTTTGNTAAACATGGACATGNTTCAGTG  
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA  
AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT  
GGGCCCAGTGTCCTCTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT  
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT  
AGTTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

Figure 1. The effect of the concentration of the *Agaricus bisporus* spores on the growth of *Agaricus bisporus* on the substrate. The concentration of the spores was 10<sup>4</sup> spores/g (a), 10<sup>5</sup> spores/g (b), 10<sup>6</sup> spores/g (c), 10<sup>7</sup> spores/g (d), 10<sup>8</sup> spores/g (e), 10<sup>9</sup> spores/g (f), 10<sup>10</sup> spores/g (g), 10<sup>11</sup> spores/g (h), 10<sup>12</sup> spores/g (i), 10<sup>13</sup> spores/g (j), 10<sup>14</sup> spores/g (k), 10<sup>15</sup> spores/g (l), 10<sup>16</sup> spores/g (m), 10<sup>17</sup> spores/g (n), 10<sup>18</sup> spores/g (o), 10<sup>19</sup> spores/g (p), 10<sup>20</sup> spores/g (q), 10<sup>21</sup> spores/g (r), 10<sup>22</sup> spores/g (s), 10<sup>23</sup> spores/g (t), 10<sup>24</sup> spores/g (u), 10<sup>25</sup> spores/g (v), 10<sup>26</sup> spores/g (w), 10<sup>27</sup> spores/g (x), 10<sup>28</sup> spores/g (y), 10<sup>29</sup> spores/g (z), 10<sup>30</sup> spores/g (aa), 10<sup>31</sup> spores/g (ab), 10<sup>32</sup> spores/g (ac), 10<sup>33</sup> spores/g (ad), 10<sup>34</sup> spores/g (ae), 10<sup>35</sup> spores/g (af), 10<sup>36</sup> spores/g (ag), 10<sup>37</sup> spores/g (ah), 10<sup>38</sup> spores/g (ai), 10<sup>39</sup> spores/g (aj), 10<sup>40</sup> spores/g (ak), 10<sup>41</sup> spores/g (al), 10<sup>42</sup> spores/g (am), 10<sup>43</sup> spores/g (an), 10<sup>44</sup> spores/g (ao), 10<sup>45</sup> spores/g (ap), 10<sup>46</sup> spores/g (aq), 10<sup>47</sup> spores/g (ar), 10<sup>48</sup> spores/g (as), 10<sup>49</sup> spores/g (at), 10<sup>50</sup> spores/g (au), 10<sup>51</sup> spores/g (av), 10<sup>52</sup> spores/g (aw), 10<sup>53</sup> spores/g (ax), 10<sup>54</sup> spores/g (ay), 10<sup>55</sup> spores/g (az), 10<sup>56</sup> spores/g (ba), 10<sup>57</sup> spores/g (bb), 10<sup>58</sup> spores/g (bc), 10<sup>59</sup> spores/g (bd), 10<sup>60</sup> spores/g (be), 10<sup>61</sup> spores/g (bf), 10<sup>62</sup> spores/g (bg), 10<sup>63</sup> spores/g (bh), 10<sup>64</sup> spores/g (bi), 10<sup>65</sup> spores/g (bj), 10<sup>66</sup> spores/g (bk), 10<sup>67</sup> spores/g (bl), 10<sup>68</sup> spores/g (bm), 10<sup>69</sup> spores/g (bn), 10<sup>70</sup> spores/g (bo), 10<sup>71</sup> spores/g (bp), 10<sup>72</sup> spores/g (bq), 10<sup>73</sup> spores/g (br), 10<sup>74</sup> spores/g (bs), 10<sup>75</sup> spores/g (bt), 10<sup>76</sup> spores/g (bu), 10<sup>77</sup> spores/g (bv), 10<sup>78</sup> spores/g (bw), 10<sup>79</sup> spores/g (bx), 10<sup>80</sup> spores/g (by), 10<sup>81</sup> spores/g (bz), 10<sup>82</sup> spores/g (ca), 10<sup>83</sup> spores/g (cb), 10<sup>84</sup> spores/g (cc), 10<sup>85</sup> spores/g (cd), 10<sup>86</sup> spores/g (ce), 10<sup>87</sup> spores/g (cf), 10<sup>88</sup> spores/g (cg), 10<sup>89</sup> spores/g (ch), 10<sup>90</sup> spores/g (ci), 10<sup>91</sup> spores/g (cj), 10<sup>92</sup> spores/g (ck), 10<sup>93</sup> spores/g (cl), 10<sup>94</sup> spores/g (cm), 10<sup>95</sup> spores/g (cn), 10<sup>96</sup> spores/g (co), 10<sup>97</sup> spores/g (cp), 10<sup>98</sup> spores/g (cq), 10<sup>99</sup> spores/g (cr), 10<sup>100</sup> spores/g (cs), 10<sup>101</sup> spores/g (ct), 10<sup>102</sup> spores/g (cu), 10<sup>103</sup> spores/g (cv), 10<sup>104</sup> spores/g (cw), 10<sup>105</sup> spores/g (cx), 10<sup>106</sup> spores/g (cy), 10<sup>107</sup> spores/g (cz), 10<sup>108</sup> spores/g (da), 10<sup>109</sup> spores/g (db), 10<sup>110</sup> spores/g (dc), 10<sup>111</sup> spores/g (dd), 10<sup>112</sup> spores/g (de), 10<sup>113</sup> spores/g (df), 10<sup>114</sup> spores/g (dg), 10<sup>115</sup> spores/g (dh), 10<sup>116</sup> spores/g (di), 10<sup>117</sup> spores/g (dj), 10<sup>118</sup> spores/g (dk), 10<sup>119</sup> spores/g (dl), 10<sup>120</sup> spores/g (dm), 10<sup>121</sup> spores/g (dn), 10<sup>122</sup> spores/g (do), 10<sup>123</sup> spores/g (dp), 10<sup>124</sup> spores/g (dq), 10<sup>125</sup> spores/g (dr), 10<sup>126</sup> spores/g (ds), 10<sup>127</sup> spores/g (dt), 10<sup>128</sup> spores/g (du), 10<sup>129</sup> spores/g (dv), 10<sup>130</sup> spores/g (dw), 10<sup>131</sup> spores/g (dx), 10<sup>132</sup> spores/g (dy), 10<sup>133</sup> spores/g (dz), 10<sup>134</sup> spores/g (ea), 10<sup>135</sup> spores/g (eb), 10<sup>136</sup> spores/g (ec), 10<sup>137</sup> spores/g (ed), 10<sup>138</sup> spores/g (ee), 10<sup>139</sup> spores/g (ef), 10<sup>140</sup> spores/g (eg), 10<sup>141</sup> spores/g (eh), 10<sup>142</sup> spores/g (ei), 10<sup>143</sup> spores/g (ej), 10<sup>144</sup> spores/g (ek), 10<sup>145</sup> spores/g (el), 10<sup>146</sup> spores/g (em), 10<sup>147</sup> spores/g (en), 10<sup>148</sup> spores/g (eo), 10<sup>149</sup> spores/g (ep), 10<sup>150</sup> spores/g (eq), 10<sup>151</sup> spores/g (er), 10<sup>152</sup> spores/g (es), 10<sup>153</sup> spores/g (et), 10<sup>154</sup> spores/g (eu), 10<sup>155</sup> spores/g (ev), 10<sup>156</sup> spores/g (ew), 10<sup>157</sup> spores/g (ex), 10<sup>158</sup> spores/g (ey), 10<sup>159</sup> spores/g (ez), 10<sup>160</sup> spores/g (fa), 10<sup>161</sup> spores/g (fb), 10<sup>162</sup> spores/g (fc), 10<sup>163</sup> spores/g (fd), 10<sup>164</sup> spores/g (fe), 10<sup>165</sup> spores/g (ff), 10<sup>166</sup> spores/g (fg), 10<sup>167</sup> spores/g (fh), 10<sup>168</sup> spores/g (fi), 10<sup>169</sup> spores/g (fj), 10<sup>170</sup> spores/g (fk), 10<sup>171</sup> spores/g (fl), 10<sup>172</sup> spores/g (fm), 10<sup>173</sup> spores/g (fn), 10<sup>174</sup> spores/g (fo), 10<sup>175</sup> spores/g (fp), 10<sup>176</sup> spores/g (fq), 10<sup>177</sup> spores/g (fr), 10<sup>178</sup> spores/g (fs), 10<sup>179</sup> spores/g (ft), 10<sup>180</sup> spores/g (fu), 10<sup>181</sup> spores/g (fv), 10<sup>182</sup> spores/g (fw), 10<sup>183</sup> spores/g (fx), 10<sup>184</sup> spores/g (fy), 10<sup>185</sup> spores/g (fz), 10<sup>186</sup> spores/g (ga), 10<sup>187</sup> spores/g (gb), 10<sup>188</sup> spores/g (gc), 10<sup>189</sup> spores/g (gd), 10<sup>190</sup> spores/g (ge), 10<sup>191</sup> spores/g (gf), 10<sup>192</sup> spores/g (gg), 10<sup>193</sup> spores/g (gh), 10<sup>194</sup> spores/g (gi), 10<sup>195</sup> spores/g (gj), 10<sup>196</sup> spores/g (gk), 10<sup>197</sup> spores/g (gl), 10<sup>198</sup> spores/g (gm), 10<sup>199</sup> spores/g (gn), 10<sup>200</sup> spores/g (go), 10<sup>201</sup> spores/g (gp), 10<sup>202</sup> spores/g (gq), 10<sup>203</sup> spores/g (gr), 10<sup>204</sup> spores/g (gs), 10<sup>205</sup> spores/g (gt), 10<sup>206</sup> spores/g (gu), 10<sup>207</sup> spores/g (gv), 10<sup>208</sup> spores/g (gw), 10<sup>209</sup> spores/g (gx), 10<sup>210</sup> spores/g (gy), 10<sup>211</sup> spores/g (gz), 10<sup>212</sup> spores/g (ha), 10<sup>213</sup> spores/g (hb), 10<sup>214</sup> spores/g (hc), 10<sup>215</sup> spores/g (hd), 10<sup>216</sup> spores/g (he), 10<sup>217</sup> spores/g (hf), 10<sup>218</sup> spores/g (hg), 10<sup>219</sup> spores/g (hh), 10<sup>220</sup> spores/g (hi), 10<sup>221</sup> spores/g (hj), 10<sup>222</sup> spores/g (hk), 10<sup>223</sup> spores/g (hl), 10<sup>224</sup> spores/g (hm), 10<sup>225</sup> spores/g (hn), 10<sup>226</sup> spores/g (ho), 10<sup>227</sup> spores/g (hp), 10<sup>228</sup> spores/g (hq), 10<sup>229</sup> spores/g (hr), 10<sup>230</sup> spores/g (hs), 10<sup>231</sup> spores/g (ht), 10<sup>232</sup> spores/g (hu), 10<sup>233</sup> spores/g (hv

## FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA  
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCCTCGGCCTCCC  
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTAAAGAAGTTAAT  
GAAACCATACCTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA  
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA  
TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA  
CTTTTAACAAAAAATTTGCATCACTTTTAAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG  
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTCGAAATCTGCCT  
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTCGA  
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC  
TTGAATATAAGACCCTACTTGCTATCTCCCCTGCAACCAGCCAGGAGCCACCCATCCTCCAGC  
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA  
TGCTCATTTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT  
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT  
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAGAGG  
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT  
AGAAGTTCAATGTGGCAAGGAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT  
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT  
GACTAGATGATAAATGCCTGTACTCCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG  
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA  
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG  
ACAATCACTTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA  
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT  
CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA  
AGACCATCCTGGCTAATACAGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG  
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA  
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG  
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

1000 900 800 700 600 500 400 300 200 100 0

## **FIGURE 33**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL  
LHLYH

**Signal peptide:**

amino acids 15-27

67  
66  
65  
64  
63  
62  
61  
60  
59  
58  
57  
56  
55  
54  
53  
52  
51  
50  
49  
48  
47  
46  
45  
44  
43  
42  
41  
40  
39  
38  
37  
36  
35  
34  
33  
32  
31  
30  
29  
28  
27  
26  
25  
24  
23  
22  
21  
20  
19  
18  
17  
16  
15  
14  
13  
12  
11  
10  
9  
8  
7  
6  
5  
4  
3  
2  
1

# FIGURE 34

GCCGCGGCGAGAGCGCGCCAGCCCCGCGCGATGCCCCGCGCGCCAGGACGCCTCCTCCCGCTGCTGGCCCCGGC  
 CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGGCGGGCGCTGGGGCGCCCGGGCCAGG  
 AGGCGGCGGCGGCGGCGGCGGACGGGCCCCCGCGGCGAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCACC  
 TGTACACGGCCGACATGTTACAGCACGGGATCCAGAGCGCCGCGCACTTCGTCATGTTCTTCGCGCCCTGGTGTG  
 GACACTGCCAGCGGCTGCAGCCGACTTGGAAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT  
 ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCCAGGGGGTGCGAGGATACCCACCTTAA  
 AGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCTCGGGACTTCCAGACACTGGAAAACTGGATGC  
 TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGGAAACCGCCAGTGCCCCGAGCTCAAGCAAG  
 GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTCTTCGCTC  
 CGTGGTGTGGTCACTGCAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCCCTGAACATTCCGAAACTG  
 TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACCTGCTCCGGAACACAGGTTCTGTGGCTATCCCACTC  
 TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGGAAAGCGGGATTTGGAGTCACTGAGGGAGTACG  
 TGGAGTCGAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCGGTGCTGGCAG  
 CTGAGCCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAAATAACTTCGATGACACCATTGCAGAAGGAA  
 TAACCTTCATCAAGTTTTATGCTCCATGGTGTGGTCACTGTGAAGTCTGGCTCCTACTTTGGGAGGAACTCTCTA  
 AAAAGGAATTCCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAACGGAATATCTGCAGCA  
 AGTATTCCGCTACGAGGCTACCCACGTTATTGCTTTTCCGAGGAGGGAAGAACTCAGTGAGCACAGTGAGGCA  
 GAGACCTTGACTCGTTACACCGCTTTGTCTGAGCCAAGCGAAAGACGAACTTTAGGAACACAGTTGGAGGTCA  
 CTCTCCTGCCCAGCTCCCGCACCCCTGCGTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCCACTGGTCTACAG  
 GTTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTCCAAGCCAACACACTCTACAG  
 ATTCTTTATTAAGTTAAGTTTTCTCTAAGTAAATGTGTAACCTCATGGTCACTGTGTAAACATTTTTAGTGCGGATA  
 TATCCCCCTTTGACCTTCTCTTGATGAAATTTACATGGTTTCCTTTGAGACTAAAATAGCGTTGAGGGAAATGAAA  
 TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC  
 CCACGAGTTCTGGAAAGGTGGCCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTCACAGTTGACTCAATAC  
 TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC  
 TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC  
 GTGCCTGAGTAAGAGAATGCTGATGCCATAACTTTATGTGTCGATACTTGTCAAATCAGTTACTGTTCAGGGGAT  
 CCTTCTGTTTCTCACGGGGTGAACATGTCTTTAGTTCTCATGTTAACACGAAGCCAGAGCCCACATGAACGTG  
 TGGATGTCTTCTTAGAAAGGGTAGGCATGGAAAATCCACGAGGCTCATTCTCAGTATCTCATTAACCTATTGA  
 AAGATTCCAGTTGTATTTGTACCTGGGGTGACAAGACCAGACAGGCTTTCCAGGCCTGGGTATCCAGGGAGGC  
 TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCCTTTTAGAGG  
 CTTGCTATACTTGGTCTGCTTCAAGGAGGTGACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC  
 CAAAGACAGATGTCAGTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTA  
 ACTCATGCTGTCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA  
 GATAGGTGTTTGTCTTTTACCATCGAGCTACTTCCATAATAACCACTTTGCATCCAACACTCTTACCCACCT  
 CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA  
 CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCACG  
 TATGGTTCACAGATAATTCTTTTAAAAAACCACCTCCTAGAGAAGCACAACCTGTCAAGAGTCTTGTACA  
 CACAACCTCAGCTTTGCATCACGAGTCTTGATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTTTTACACTAT  
 GATACTTTCTAAATAAACTCTTTTAAAAA

100229533660

[illegible]

><subunit 1 of 1, 432 aa, 1 stop

MPARPGRLLPLLARPAALTALLLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS  
KHLYTADMFTTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH  
SDVCSAQGVRGYPTLKLKFKPGQEAVKYQGPRDFQTLLENWMLQTLNEEPVTPEPEVEPPSAPE  
LKQGLYELSASNFEHLHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT  
QHYELCSGNQVRGYPTLLWFRDGGKVDQYKGKRDLESLREYVESQLQRTETGATETVTPSEA  
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA  
GVKIAEVDCTAERNICSKYSVRGYPTLLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

amino acids 1-32

## FIGURE 36

CTTTTCTGAGGAACCAACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT  
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA  
CCGCTGAAGTCTGTGCCACACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA  
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA  
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTGGGAAGA  
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT  
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACTGGATATTAGTATTGCTCGGCT  
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAACTGAAGAGAAAT  
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG  
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC  
CAAGAGTGGCTTCTTTTCGGGTGTTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA  
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC  
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG  
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT  
ACGTATTTGCTATTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG  
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT  
CCATCATCAAAAAAAAAAAAAAAAAA

\_\_\_\_\_

><MW: 30645, pI: 7.47, NX(S/T): 2

amino acids 1-25

[illegible]

## FIGURE 38

GGTTCTATCGATTCTGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC  
CCACGCGTCCGCTGCTCTCCGCCCCTGTGGAGTGGTGGGGGCCTGGGTGGGAATGGGCGTGT  
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTTT  
CGGAAGGGAGGATCAGGGATGTTTGCAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC  
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA  
AGCTACACCTCTGGCCGCGATTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT  
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG  
TCCCAGAGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACCTGGCCCAGCAGCGCGCCGCGC  
ACACCTTTCTCATTCACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC  
AGGGCTGCACGCGCCTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG  
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG  
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT  
CTGTACCTGGAGCAACTGTGGCGCTGCTCCTCCCGCTGGCCCAGAGTTTCTGTGGCTCTG  
GTTCCGGGCTGGCCAAGGCCGCTGCGCACTGCCTTTGTGCCACCCGCCCTGCGCCGGGGCC  
CCCTGCTGCACTGCCTCCGCGAGCTGCGGCGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTCTG  
GAGTCCCTGGAGCCGGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG  
CCCAGGAACCCACCCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG  
GGCCAGTGCCAGGATACCTCTCTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC  
ACCTCTGGCACCACGGGCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG  
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC  
TCTACCACATGTCCGGTTCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG  
GTGCTGAAATCCAAGTTCTCGGCTGGTCACTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC  
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCGAGCAAGGCAG  
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATAACCTGGGAGCGT  
TTTGTGCGGCGCTTCGGGCCCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT  
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCCTGGCTTTACAAGC  
ATATCTTCCCCTTCTCCTTGATTGCTATGATGTCACCACAGGAGAGCCAATTCGGGACCCC  
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA  
GCAGTCCCCATTCTTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG  
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA  
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC  
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG  
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC  
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCC  
GCCCCGATTCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG  
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCTGTCTGACCCACTGTACGTTCTGGAC  
CAGGCTGTAGGTGCCTACCTGCCCCCTCACAACCTGCCCGGTACAGCGCCCTCCTGGCAGGAAA  
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGG  
CCGTTGCAGGTGTACTGGGCTGTGAGGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT  
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC  
TTGGCCGCCATGGCCCAACTTGTTTATTGCAG

594553 9499

## **FIGURE 39**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTRAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLPLLLLLPL  
LLLKLHLWPQLRWLPADLAFVRLCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ  
RAAHTFLIHGSRFSYSEAERESNRAARAFRLALGWDWGPDPGGDSGEGSAGEGERAAPGAGD  
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGGLAKAGLRTAFVPTAL  
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA  
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL  
ALPLYHMSGSLLGIVGCMGIGATTVLKSFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP  
SKAERGHKVRRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW  
LYKHIFPFSLIRYDVTTGEPIRD PQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQQGK  
LLKDVFRPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV  
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK  
QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

### **Type II transmembrane domain:**

amino acids 45-65

### **Other transmembrane domain:**

amino acids 379-398

### **cAMP- and cGMP-dependent protein kinase phosphorylation site**

starting at amino acid 136

### **CUB domain protein motif**

amino acids 254-261

### **putative AMP-binding domain signature**

amino acids 332-343

### **N-glycosylation sites**

amino acids 37-40 and 483-486

## FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA  
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATAACACTCGCTCTC  
GCTTGTCCATCTCCCTCCCGGGGAGCCGGCGCGCTCCACCTTTGCCGCACACTCCGGC  
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC  
CCATGGTGGTTTTTTTAAACACTTCTTTTCTTCTCTTCTCCTCGTTTTTGATTGCACCGTTTCCA  
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTGGCTTGCCATCGT  
CCATCTGGCTTATAAAAGTTTGTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG  
CTGGCAGAAGGGGGTGACGCTGGGCAGCGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT  
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG  
TTGCACCA**ATG**CCTTCTTGGATCGGGGCTGTGATTCTTCCCCCTCTTGGGGCTGCTGCTCTCCC  
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC  
AAGGGATTACAGCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG  
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC  
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG  
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA  
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTGAGAAGTCTTCCAGGACCTCT  
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAAATGTGAATCTGGAGGAAATGCTCAATGAC  
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCTCAGTATCACTTCAGTGA  
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC  
GGAAACTGAAGATTACAGGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG  
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTGAGCCCAACCCAGGGTGTATCCG  
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTGCGGGGCTTCCCCTGTGAGGCCCTGCA  
ACAACTACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG  
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT  
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA  
ACAGCATGCAGGTGTCTGCAAAGGTCTTTCAGGGATGTGGTCAGCCCAAACCTGCTCCAGCC  
CTCAGATCTGCCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTTCAGGCCCTACAATCCTGA  
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA  
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC  
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT  
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA  
TCACTCGGCCTGACACTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA  
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG  
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTTTGTACCA  
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC  
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTTGGCACTGCAGAGACTGTGCAGAT**A**  
**A**TCTTGGGTTTTTGGTCAGATGAAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT  
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACTTACCGTTTTCTATGAGAAGAGAG  
CAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACCGGGTGCCAGACTGAACTG  
CTTCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA  
ATTTTTCGTACCAGGAGATTTTCTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT  
CTCACGTTGTGAGGGTTTTTTTTTCTCATTAAAT

## FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADV KARS CGEVRQAYGAKGFS LADIPYQEIAGEHLRICPQ  
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMF  
VRTYGMLYMQNSEVFQDLFTTELKRYYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY  
LECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL  
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES  
VMDPIDVKISEAIMNMQENSMQVS AKVFQCGQPKPAPALRSARSAPENFNTRFRPYNPEER  
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDES VTAGTSNEEECWNGH SKARYLPE  
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG  
SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

## FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG  
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG  
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT  
GGAGGGAGATCAGGAAACGGCTTCTTCTCACTTCGCCGCTGGTGAGTGTGCGGGAGATTGGCAAACGCCTAGG  
AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGAGAAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG  
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA  
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA  
AAAGTTCCTGTCCACTGTGATTCTCAATTCTTGTCTTGGTTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA  
ACTTTTTTCTTTTTTTTTTTTCTTGGTGAAAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAAAGTGAAATGTGC  
TGGAGAAGAGCGAGCCCTCCTTGTCTTCCGGAGTCCCATCCATTAAAGCCATCACTTCTGGAAGATTAAAGTTGT  
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTACCCTGTGTGGGTGCATG  
TGTGCGCCCGCAGCGGCGCGGGGCGCGTGGTTCTCCGCGTGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA  
GGGGCTGTGCGGGGCATCCGCCTCCGCCTTCTCCACAGGCCTGTGTCTGTCTTGAAAGATGCTAGCAATGGGGG  
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTTGGGGCCAGGCCTTAGAAGAGGAGG  
AAGAAGGGGCCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCA  
TTTTCATCTAGCGGATGATCAGGGATTTAGAGATGTGGGTACCACGGATCTGAGATTAAAAACCTACTCTTG  
ACAAGCTCGCTGCCGAAGGAGTTAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT  
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTTTAC  
CTCTGGACAATGCCACCCTACCTCAGAACTGAAGGAGGTTGGATATTCAACGCATATGGTCGGAAAATGGCACT  
TGGGTTTTTAACAGAAAAGAATGCATGCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG  
GGGATTACTATACACACTACAAATGTGACAGTCTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG  
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCATA  
ACCCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCACTTACCCTGCAAGCTCCTGGCAGGTATT  
TCGAACACTACCGATCCATTATCAACATAAAACAGGAGAAGATATGCTGCCATGCTTTCTGCTTAGATGAAGCAA  
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG  
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC  
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTGTGACATCACTG  
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT  
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGCATAACATTGACCCCTATACACCAAGGC  
AAAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG  
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCCTCAGTCTTTCAGCAACCTGGGACCGAACCG  
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA  
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAAC  
TGCAGTGCCGGTCAGGTATCCCCCAAGACCCCAAGTAACCCTAGGCTCAATGGAGGGGTCTGGGGACCATG  
GTATAAAGAGGAAACCAAGAAAAAGGCCAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAGCAAAAAAA  
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA  
GGCTTGTTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCACACTGAAAACGTCTCTGCTCAGTG  
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG  
GTGAGTCCTGAGTTCCACTGCTGTGCTTCAGTCAACTGACCAAACACTGCTTTGAATTATAGGAGGAGAAACAATA  
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAAAACTACCTTTGATAAATTAC

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG

[illegible][illegible]

## **FIGURE 43**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296

><subunit 1 of 1, 515 aa, 1 stop

><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEEGALLAQA  
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHGSEIKTPTLTDKLAAEGVKLENYYVQPICTP  
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMGKWHLGFNRKEC  
MPTRRGFDFTFFGSLGSGDYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ  
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA  
LKTYGFYNNSIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGVHSPLLKNKGTVCK  
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG  
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS  
TSQPTHMRGWTYLTGIQES

### **Important Features:**

#### **Signal Peptide:**

amino acids 1-37

#### **Sulfatases signature 1.**

amino acids 120-132

#### **Sulfatases signature 2.**

amino acids 168-177

#### **Tyrosine kinase phosphorylation site.**

amino acids 163-169

#### **N-glycosylation sites.**

amino acids 157-160, 306-309 and 318-321

## FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC  
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC  
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG  
GTGGTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTGAGCCT  
GGGGTCTGTCACTATGGAACATAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG  
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAACAAAT  
GCAGATGCTTTCCAGGATACACCGGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG  
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT  
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACTCTAGGACATGTGCCATGATAA  
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA  
CTCCGCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT  
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTGACA  
TTGGTTTCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT  
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCTTCAAGTG  
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTTCTGCTATCCCTGAAAATTCTG  
TGAAGGAAGTCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC  
AAAAACAGCATGAAAAAGAAGGCCAAAATTAATAATGTTACCCCAAGACCCACCAGGACTCC  
TACCCCTAAGGTGAACTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAT  
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAAAGAAAGAGGGGCTTGAGGATGAGAAAAGAG  
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT  
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA  
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT  
GGAAACAGGATAGAGAAGATGATTTTGAAGTGAATCCTGCTGATCGAGATAATGCTATTGGC  
TTCTATATGGCAGTTCGGCCTTGGCAGGTGACAAGAAAGACATTGGCCGATTGAACTTCT  
CCTACCTGACCTGCAACCCCAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG  
ACAAAGTCGGGAACTTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG  
ACCACGAGTGAGGATGAAAAGTGGAAGACAGGGAAAAATTGAGTTGTATCAAGGAACTGATGC  
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG  
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG  
TACTATCTTTATATTTGACTTTGTATGTCAGTTCCCTGGTTTTTTTGATATTGCATCATAG  
GACCTCTGGCATTTTAGAACTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG  
TAAGATGCCTTTCTGTATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT  
TCTCAGTCATTTCTGAATCTTTCCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC  
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA  
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGGA  
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC  
TTGTATATTTAATTCTTTGTAATAATAA

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC

## **FIGURE 45**

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNRNSKGVCE  
ATCEPGCKFGECVGPNNKCRFCFPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSKYKCFCLSGH  
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP  
YNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK  
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPK  
VNLQPFNYEEIVSRGNSHGGKKGNEEK

### **Signal peptide:**

amino acids 1-21

### **EGF-like domain cysteine pattern signature.**

amino acids 80-91

### **Calcium-binding EGF-like domains**

amino acids 103-124, 230-251 and 185-206

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG  
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC  
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG  
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTCAGCCAGAAGAGTGCAT  
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT  
TGTTTTGCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC  
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCCAGCGTTCTCTGTGC  
ATGGATACCAGCTTGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC  
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA  
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT  
CTCCGGGGTTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC  
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA  
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG  
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT  
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG  
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAATC  
TTTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG  
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT  
ACTTTTAAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAAG  
ATTGCCATGAATCTTGCAAAA

1000  
900  
800  
700  
600  
500  
400  
300  
200  
100  
0

## **FIGURE 47**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343  
><subunit 1 of 1, 289 aa, 1 stop  
><MW: 32268, pI: 9.21, NX(S/T): 0  
MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPDL  
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL  
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP  
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW  
QYMPTWAWWITNKMGGKKRIENFKSGVDADSSYFKIFKTKHD

### **Important Features:**

#### **Signal Peptide:**

amino acids 1-31

#### **Transmembrane domain:**

amino acids 136-157

#### **Tyrosine kinase phosphorylation site.**

106-113 and 107-114

#### **Homologous region to Short-chain alcohol dehydrogenase**

amino acids 80-90, 131-168, 1-13 and 176-185

[illegible]

CGACAGCTGGGCACCGCCATCAGCTGTTTCGCGCGCTCTTCTCCTCCAGGTGGGGCAGGGGTTTC  
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTTCG  
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC  
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTT  
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT  
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA  
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGT  
GACAGCGTGGCCCCGGCCCCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCC  
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC  
CTGGGTCCCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT  
CAGTGGTCTCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG  
GCTGCCCTGCAGTTGCCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT  
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC  
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC  
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACCTGTATCTACAACCAGCT  
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCCTGGGATGCTATGTGGGGGGCCCCCAGCCTG  
GGGTGCAGGGCCCCCTGTCAGGGAGATTCCGGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA  
CACTGGGTTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT  
GCTGCTGACCAACACAGCTGCTCACAGTTCCTGGCTGCAGGCTCGAGTTTCAGGGGGCAGCTT  
TCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA  
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCTCCTCCCATGGCCCTGGGAGGCCAGGCT  
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG  
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA  
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG  
CTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC  
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG  
GCCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG  
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG  
TGTGTACCAAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGTCTGGGGCACCCTGGTG  
CATGAGGTGAGGGGCACATGGTTCTTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAG  
CCCCGCCAGGCCGGCGGTCTTACC CGCCTCCTGCCTATGAGGACTGGGTGAGCAGTTTGG  
ACTGGCAGGTCTACTTCGCCGAGGAACAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC  
AACATAAGCCAACCAACCAGCTGCTGAGCAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC  
AGGCAGGCAAAATGGCATTACTGCCCTGTCTCCTCCACCCCTGTATGTGTGATTCCAGGCAC  
CAGGGCAGGCCCAGAAGCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA  
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACTCTGCTACCAAGC  
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTCAGATACAATCACGCCAGCCACGTTGTTT  
TGAAAATTTCTTTTTTTTGGGGGGCAGCAGTTTTTCCTTTTTTTTAAACTTAAATAAATTGTTAC  
AAAATAAAA

**SECRET**

MLLSSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVVQEPQGKAKRHGNTV  
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA  
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS  
DAPGTLRLNRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVLC  
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS  
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV  
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLPDGERG  
WVLGRARPGAGISSLQTVPVTLGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS  
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEF  
GSCANISQPTSC

Signal peptide:

Homologous region to Serine proteases, trypsin family

N-glycosylation sites.

## Kringle domains

amino acids 79-96, 343-360 and 235-247

## FIGURE 50

CGGGCCGCCCCCGGCCCCATTTCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG  
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG  
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT  
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG  
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC  
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA  
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC  
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG  
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTTCGGCAGCGCTGCCGTGAGCCT  
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC  
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT  
TGGCAGGATGGGCAGGGTGTGCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGA  
GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGCTGGGTGCGAATGGCACCTACA  
GCTGCCTGGTGCACAAACCCGTCGTGCAGCAGGATGCGCACRGCTCTGTACCATCACAGGG  
CAGCCTATGACATTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT  
TGCACTGCTGGTGGCCCTGGCTTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG  
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT  
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTTGACCATGAGGACCAGG  
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC  
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC  
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT  
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT  
GCCTTATTTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCTC  
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAGATGCAATATTGAGA  
CTGACTGACCCCTGCCTTATTTACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTC  
TCCAATGGCCGTGATACACTAGTGATCATGTTTACGCCCTGCTTCCACCTGCATAGAATCTTT  
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCTTTC  
CCCTCCTTCCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG  
AGGGGACTGCCCCCACCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCCTGGC  
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG  
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC  
TCGGAGGGATTTTGTAACCTGGGGGTATATTTTGGGGAAAATAAATGTCCTTTGTAAAAAAA  
AAAAAAAAAAAAA

## **FIGURE 51**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ  
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVVADEGSFTCFV  
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITCSSYQGYPEAEVFWQDGQGVPL  
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVARNPVLQQDAHXSVTITGQPMTFPPEA  
LWVTVGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGEGEKSKTALQPLKHSDSKED  
DGQEIA

### **Important features:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 251-270

#### **N-glycosylation site.**

amino acids 91-94, 104-107, 189-192 and 215-218

#### **Homologous region to Immunoglobulins and MHC**

amino acids 217-234

## FIGURE 52

TTCGTGACCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC  
CTGAACCTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACACGTTCTTAAATCTATGAAGTCG  
AGGGACCTTTTCGCTGCTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAACATGAGGCTTTTCT  
TGTGGAACGCGGTCTTGACTCTGTTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA  
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCAAAGGAGGGGA  
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC  
ACAAACATAACAATGGTCAGCCCATTGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT  
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGC  
TCTGGGCTATGGAAGAAGGAAAAGGTAAAATTCCCCAGAAAGTACACTGATATTTAATA  
TTGATCTCCTGGAGATTCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT  
AATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA  
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGTATAAAG  
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA  
TAGAGATACATCTACCCCTTTTAAATATAGCACTCATCTTCAAGAGAGGGCAGTCATCTTTAA  
AGAACATTTTATTTTTATACAATGTTCTTTCTTGCTTTGTTTTTATTTTTATATATTTTTT  
CTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT  
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTACAG  
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAATGTTGCAACTGGGAATATACC  
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCCTATATTTCTGCTTCCCTCTATTTTC  
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT  
GTTATAATGAAATAGTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA  
TGTTTGTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG  
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG  
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA  
CCCTATCTCTACTAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCCTGTAATCCCAGCT  
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG  
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA  
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT  
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG  
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC  
TAGCGGAATATCCTTCCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA  
TTGTATCATAAGATAAAGTAGTAAACAGTCTACATTTTCCCATTCTGTCTCATCAAAAC  
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG  
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAAACATGGTGAAACCTTGTCTCTA  
CTAAAAATACAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG  
GCTGAGACAGGAGATTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC  
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA  
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

CCGCGGCGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA

## **FIGURE 53**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTFLFVTSLLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSL
FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVED
IFDKEDDKDGFISAREFTYKHDEL
```

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **N-glycosylation site.**

amino acids 176-179

#### **Casein kinase II phosphorylation site.**

amino acids 143-146, 156-159, 178-181 and 200-203

#### **Endoplasmic reticulum targeting sequence.**

amino acids 208-211

#### **FKBP-type peptidyl-prolyl cis-trans isomerase**

amino acids 78-114 and 118-131

#### **EF-hand calcium-binding domain.**

amino acids 191-203, 184-203 and 140-159

#### **S-100/ICaBP type calcium binding domain**

amino acids 183-203

[illegible]

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTTAGCAT  
 CCAACCATTCTCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCCTCTCCCGTAGCCCACCCGA  
 CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT  
 CACGGGGCTCAGTCTCTTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC  
 CTGCCACCCTCAACGTCTCAATGGCTCTGACGCCCCGCTGCCCTGCACCTTCAACTCCTGC  
 TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTC  
 TGAGGAGATGTTCTCCAGTTCCGCATGAAGATCATTAACTGAAGCTGGAGCGGTTTCAAG  
 ACCGCGTGAGATTCTCAGGGAACCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTG  
 CAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCCTGACCGCCACCGTGG  
 CCATGGCAAGATCCATCTGCAGGTCCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG  
 CCGTGATTGTGGGTGCCCTCCGTGCGGGGCTTCTTGCTGTGGTTCATCTTGCTGCTGATGGTG  
 GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA  
 GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC  
 CTGCAGCCTCCCCTGTCCCCTCTCTCCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCG  
 CTCTTGCTGTGCTTCCCCTGACCTAGGACCCAGGGGCCACCTGGGGCCTCCTGAACCCCCG  
 ACTTCGTATCTCCACCCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA  
 TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGG  
 GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA  
 GGAGGGGGCCGCTGTACCTGCCAGTGCTTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGG  
 GGAGGGAGGGCTTTCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG  
 TGCTCCTCCCCTGCTCCCAGCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA  
 AACTTGAGGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG  
 CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGGCCTCGTTTCCATT  
 GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG  
 ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC  
 AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG  
 GCTCTGCCTTCTCCATGGGGTAACCACCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT  
 GAGGAGGCCATGCACAGGGTGGGGCAGCTTTCTTTGGGGCTTCAGTGAGAACTCTCCCAGTT  
 GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGGAAGGGAAGCCTGAGGCCG  
 GCATAAGGGGAGGCCTTGGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG  
 CTACTCGCTCCTCTCCCAACAACCTCCCTTCGTGGGGACAAAGTGACAATTGTAGGCCAGGC  
 ACAGTGGCTCAGCGCTGTAATCCCAGCACTTTGGGAGGCCAAGCGGGGTGGATTACCTCCAT  
 CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG  
 GTGGCGTGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCCG  
 GGAAGCAGAGGTTGCAGTGAACCTGAGATAGTGATAGTGCCACTGCAATTACAGCCTGGGTGAC  
 ATAGAGAGACTCCATCTCAAAAAAAA

## **FIGURE 55**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ  
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRVQPEDEGI  
YNCYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK  
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domain:**

amino acids 161-179

#### **Immunoglobulin-like fold:**

amino acids 83-127

#### **N-glycosylation sites.**

amino acids 42-45, 66-69 and 74-77

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTTTAGCATCCAACCATCCTCCCTTGTA  
 GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA  
 AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT  
 TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCCTCAACGTCC  
 TCAATGGCTCTGACGCCCCGCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC  
 AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCCCTCCAG  
 TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGAGGTTCTCAGG  
 GAACCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA  
 TTTACAACCTGCTACATCATGAACCCCCC

## **FIGURE 57**

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCCACCAGGACGGAGCATGGAGGTNCACATA  
CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCCGCCTGCCCTGCACCTTCAACTCCNG  
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC  
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG  
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGC  
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCC

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCCACCAGGACGGAGCATGGAGGTNCACATA  
CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCCGCCTGCCCTGCACCTTCAACTCCNG  
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC  
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG  
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGC  
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCC

## FIGURE 58

TGCGGCGACCGTCGTACACCATGCGGCCTCCACCTCCGCCCCCTACCGTGTGGGGCTGCTCCCCGGATGGCCTCCTGT  
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGGACGTACCCCCCAGTGGTGCTGGTCC  
CTGGTGATTTGGGTAAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACTACCTCTGCTCCAAGAAGA  
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACCTGCTGCTGCCTGTCATCATTGACTGCTGGATTGACAATA  
TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG  
GGAAGACCTTCTCACTGGAGTTCCTGGACCCAGCAAAAGCAGCGTGGGTTCCTATTTCCACACCATGGTGGAGA  
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG  
AAAACGGGCCCCCTACTTCCTGGCCCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCGTGGTGC  
TGCTTGGCCACAGTATGGGCAACATGTACACGCTCTACTTTCTGTCAGCGGCAGCCGAGGCCTGGAAGGACAAGT  
ATATCCGGGCTTCGTGTCACTGGGTGCGCCCTGGGGGGCGTGGCCAAGACCCTGCGCGTCCCTGGCTTCAGGAG  
ACAACAACCGGATCCCACTCATCGGGCCCCCTGAAGATCCGGGAGCAGCAGCGGTGAGTGTCTCCACCAGCTGGC  
TGCTGCCCTACAACCTACACATGGTCACCTGAGAAGGTGTTCTGTGACACACCCACAATCAACTACACACTGCGGG  
ACTACCGCAAGTTCCTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG  
AAGCCACGATGCCACCTGGCGTGCAGCTGCCTCTATGGCTACTGGCGTCCCAACACAGACTCCTTCTACT  
ATGAGAGCTTCCCTGACCGTGACCTTAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACCTGAAGAGTGGCC  
TGCAGTGCCAGGCCTGGCAGAGCCGCCAGGAGCACCAAGTGTTGCTGTCAGGAGCTGCCAGGCAGCGAGCACATCG  
AGATGCTGGCCAACGCCACCACCCTGGCCTATCTGAAACGTGTGCTCCTTGGGGCCCTGATCCTGTGCCACAGGA  
CTCCTGTGGCTCGGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTGATGGCCACGCGTTTTGCAAAGTTTGTGA  
CTCACCATTCAGGCCCCGAGTCTTGAGCTGTGAAGCATCTGCCATGGGGAAGTGTGTTTGTATCCTTTCTCT  
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGGA  
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGGCCCTGGTCCAGTCCCTGCCCTGGGGCCATG  
TGTCCCCCTATTCTGTGGGCTTTTCTACTTGCCTACTGGGGCCCTGGCCCCGAGCCTTCCCTATGAGGGATGTT  
ACTGGGCTGTGGTCTGTACCCAGAGGTCCCAAGGATCGGCTCCTGGCCCCCTCGGGTGACCTTCCACACACCA  
GCCACAGATAGGCCTGCCACTGGTCTATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC  
TGACTGGCTTCCCTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGCAGTTTGTGCGTTCCTCGTGGTTCACAGGC  
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAAGCTCTGGATTGGGCAGCAGATGTG  
CCCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCTCGGATGTGCTATTTGGCCCCAGGACTGAAGCTGC  
CTCCCTTCAACCTGGGACTGTGGTTCCAAGGATGAGAGCAGGGGTGGAGCCATGGCCTTCTGGGAACCTATGGA  
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCACCATCA  
CACTGCCACCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCACAGGGCTGAGGATGGGGCTCCTATCCAC  
CCTGGCCAGCACCCAGCTTAGTGCTGGGACTAGCCAGAACTTGAATGGGACCCTGAGAGAGCCAGGGGTCCCC  
TGAGGCCCCCTAGGGGCTTTCTGTCTGCCCCAGGGTGCTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT  
CAGGGCTGCCTTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGT  
GGGTTCCTCAAAGACGCCTTCAGGCTGGACTGAGTGTCTCTCCACAGGGTTTCTGTGTCAGCTGGATTCTCTG  
TTGCATACATGCCTGGCATCTGTCTCCCTTGTCTCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG  
GATCTCGCAATAAAAGTACTCTGGATGCTGTAAAAA

CGCGGACCGTCGTACACCATGCGGCCTCCACCTCCGCCCCCTACCGTGTGGGGCTGCTCCCCGGATGGCCTCCTGT

## **FIGURE 59**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH  
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPGFGKTFSL  
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ  
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI  
PVIGPLKIREQQRSVSTSWLLPYNYSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLM  
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ  
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG

**Important features:**

**Signal peptide:**

amino acids 1-28

**Potential lipid substrate binding site:**

amino acids 147-164

**N-glycosylation sites.**

amino acids 99-102, 273-276, 289-292 and 398-401

**Lipases, serine proteins**

amino acids 189-201

**Beta-transducin family Trp-Asp repeat**

amino acids 353-365

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG  
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA  
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCCTGCATCTATGGTG  
AGGGCTACAGCAATGCCACGAGTCTAAGCAGATGTAAGCGTGTTCAACCGCAACGAGGAT  
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT  
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG  
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC  
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGGCAGCCAT  
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT  
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT  
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACCTACCAACAGCCACCCTTCACCCAGAA  
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTGAGTGGCGGTTAGCGTGGGAA  
GGGGGACAGAGAGGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAACCTGCCA  
GCCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC  
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCAGAGGGCTTCAGTCAGCCGCT  
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCCTCGCTTTTAATGA  
CCTCAGCCCCGCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT  
CAGCTTCCCCCGGCCGGGTGAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG  
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCTCA  
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT  
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG  
CTCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAAATATTC  
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCCAGACTCTGTC  
TGTGCCGAGTGTATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTGGAGACG  
GAATAAATGTTTTCTCATTCAAAG

## **FIGURE 61**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEYGSNAHESKQMYCVFN  
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIGDLLFSALWTFLWFGFC  
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP  
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

### **Important features:**

#### **Type II Transmembrane domain:**

amino acids 1-45

#### **Other transmembrane domains:**

amino acids 74-90, 108-126 and 145-161

#### **N-glycosylation site.**

amino acids 97-100

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

## FIGURE 62

GAGCCACCTACCCTGCTCCGAGGCCAGGCCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGG**ATG**  
 CCCGTGGCCGAGGCCCCCCAGGTGGCTGGCGGGCAGGGGGACGGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG  
 ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCCCGGGGCTACCTCCGCCTGGTGCCCTGTTGTGCTGCTG  
 GCCCTGCTCGTGCTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCCCTAGGGTACAAGGCGGAGGTGATGGTCAGC  
 CAGGTGTACTCAGGCAGTCTGCGTGTACTCAATCGCCACTTCTCCCAGGATCTTACCCGCCGGGAATCTAGTGCC  
 TTCCGCAGTGAAACCGCCAAAGCCCAGAAGATGCTCAAGGAGCTCATCACCAGCACCCGCTGGGAACCTACTAC  
 AACTCCAGCTCCGTCTATTCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTTCATTCTCCAAATCCCCGAG  
 CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGAGGCACTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGC  
 TCGGCTGCCGTCCCCTACAGGGCCGAGTACGAAGTGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA  
 GACATAGCTGCATTGAATTCCACGCTGGGTTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCCGGCTG  
 AAGGGGCCCTGACCACCTGGCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGACCTCATGCTCAAACCTCCGG  
 CTGGAGTGGACGCTGGCAGAGTGGCGGGACCGACTGGCCATGTATGACGTGGCCGGGGCCCTGGAGAAGAGGCTC  
 ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTTCTGGCGTCGGGGGCCATCATGGCGGT  
 GTCTGGAAGAAGGGCCTGCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCGT  
 GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGCACCCCGTACTTCCCCAGCTACTAC  
 TCGCCCCAAACCCACTGCTCCTGGCACCTCACGGTGGCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGATGCC  
 TATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGT  
 GGCTTGCGCATCCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACC  
 TCCAGATCTCCCTCACCGGGCCCCGGTGTGCGGGTGCACTATGGCTTGTAACAACAGTCGGAGCCCTGCCCTGGA  
 GAGTTCCTCTGTCTGTGAATGGACTCTGTGTCCTGCTGCTGATGGGGTCAAGGACTGCCCAACGGCCTGGAT  
 GAGAGAACTCGCTTTGCAGAGCCACATTCCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCCAAGGTCTGT  
 GATGGGCAGCCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCC  
 TTCCAGTGTGAGGACCGAGCTGCGTGAAGAAGCCCCAACCCGAGTGTGATGGGCGGCCGACTGCAGGGACGGC  
 TCGGATGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGATTGTTGGTGGAGCTGTGTCTCCGAG  
 GGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGGTCGACACATCTGTGGGGGGGCCCTCATCGCTGACCGC  
 TGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCTCCACGGTGTGTGGACCGTGTCTCCGGG  
 AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCA  
 GAAGAGGACAGCCATGACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGGCTCGGCCGCCGTGCGC  
 CCCGTCTGCCCTGCCCCGCGCTCCCACTTCTTCGAGCCCGCCTGCACTGCTGGATTACGGGCTGGGGCGCCTTG  
 CGCGAGGGCGGCCCATCAGCAACGCTCTGCAGAAAGTGATGTGCACTTGCATCCACAGGACCTGTGCAGCGAG  
 GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGT  
 GACTCAGGTGGTCCGTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTG  
 GGCTGTGGCCGGCCTAACTACTTCCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTG  
 ACCT**GAG**GAACTGCCCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCCAGGGCAACTGCCAAGCAGG  
 GGGACAAGTATTCTGGCGGGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTTGTCTCGTCC  
 CTGATGTCTGCTCCAGTGATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAAGACGTCCCCTGAGGACC  
 CAGGCCACACCCAGCCCTTCTGCCCTCCCAATTCTCTCTCCTCCGTCCCCTTCCCTCCACTGCTGCCTAATGCAAG  
 GCAGTGGCTCAGCAGCAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG  
 CTGTTTGGGCAGCCTTGCCCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCCCTGGTCTAACTTGGGATCTGGGAAT  
 GGAAGGTGCTCCCATCGGAGGGGACCTCAGAGCCCTGGAGACTGCCAGGTGGGCTGCTGCCACTGTAAGCCAA  
 AAGGTGGGGAAGTCTGACTCCAGGGTCTTGCCCCACCCCTGCCTGCCACCTGGGGCCCTCACAGCCCAGACCT  
 CACTGGGAGGTGAGCTCAGCTGCCCTTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAAAA

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

## FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152

><subunit 1 of 1, 802 aa, 1 stop

><MW: 88846, pI: 6.41, NX(S/T): 7

MPVAEAPQVAGGQGDGGDGEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL  
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT  
YNNSSSVYSFGEGPLTCFFWFILQIPEHRRMLLSPEVVQALLVEELLSTVNSSAAVPYRAEY  
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGQVLRLKGPDHCLASSCLWHLQGPKDLML  
KLRLEWTLAECRDLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY  
YDPFVLSVQPVVFQACEVNLTLDNRLDSQGVNSTPYFSPSYSPQTHCSWHLTVPSLDYGLAL  
WFDAYALRRQKYDLPTQGWITQNRRLCGLRILQPYAERIPVATAGITINFTSQISLTGP  
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS  
LPKVCQGPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPDGRDGSDEEHCD  
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITAACHCFQEDSMASSTVL  
WTVFLGKVVQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA  
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG  
KKDACQGDSSGGLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIIQQVVT

### Important features:

#### Type II transmembrane domain:

amino acids 46-67

#### Serine proteases, trypsin family, histidine active site.

amino acids 604-609

#### N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447  
and 509-512

#### Kringle domains.

amino acids 746-758 and 592-609

#### Homologous region to Kallikrein Light Chain:

amino acids 568-779

#### Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

[illegible]

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC  
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT  
CTCCCTCACCGGGCCCGGTGTGCGGGTGCACCTATGGCTTGTACAACCAGTCGGACCCCTGCC  
CTGGAGAGTTCTCTGTTCGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC  
TGCCCCAACGGCCTGGATGAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA  
CAGCACATGCATCTCACTGCCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG  
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG  
AGCTGCGTGGAAGAAGCCCCAACCCGCAGTGTGATGGGCGGCCCCGACTGCAGGGACGGCTCGGA  
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT  
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCTGGGGTGCACACATCTGTGGG  
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCCACTGCTTCCAGGAGGACAGCAT  
GGCCTCCACGGTGCTGTGGACCGTGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG  
GAGAGGTGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT  
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTTCGCTCGGCCCGCGTGCGCC  
CGTCTGCCTGCCC GCGCTCCCACTTCTTCGAGCCCGCCTGCACTGCTGGATTACGGGGCT  
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG  
ATCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC  
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGTGACTCAGGTGGTCCGCTGGTGTGCA  
AGGCACTCAGTGGCCGCTGGTTCTCTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG  
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT  
GACCTGAGGAACTGCCCCCCTGCAAAGCAGGGCCCCACCTCCTGGACTCAGAGAGCCCAGGGC  
AACTGCCAAGCAGGGGGACAAGTAT

[illegible][illegible]

## **FIGURE 66**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE  
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL  
PLAHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV  
GRITGLDPAGPMFEGADIAHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF  
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLQNQDKPSFAFQCTDSNRFKKGICLS  
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

### **Important features:**

#### **Signal peptide:**

amino acids 1-16

#### **Lipases, serine active site.**

amino acids 163-172

#### **N-glycosylation sites.**

amino acids 80-83 and 136-139

## FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCGGGGCGAGCCACCTCTTCCCCCTCCCCGC  
TTCCCTGTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT  
CGGCAAAGTTTGGCCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGCGCTCG  
CTGCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG  
AGAAGAGTGC GGCGGCGGACGGAGAAAACAACCTCCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCG  
CCGCCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTTCGAGCCCCGGGCCATGGAGCCCCCTGGGGAGGCGG  
CACCAGGGAGCCTGGGCGCCCCGGGCTCCGCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTTCCG  
GCACCTCTGGACAGCCCAGGATGCTGTTGGCCACCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG  
ACCGGATTATTTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTTAGAAGTGCAGGGCACCTTACAGA  
GGCCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAACAGACTG  
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCTACGCTCCCCCTCTCCAGCCAC  
TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCGGGGGCAACGTCAACCATCACTTACAGCTATG  
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG  
AGTTTCAGTGCCTGAACACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT  
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCCCCTCCCTGCCTTGCAATG  
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCCTGGATATACACACCTAGCCTCAGTCTCCACCCCCAGT  
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGGCTGGCCGTGCGCTTCACAGCCCTGGACTTGGGCTTTG  
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCACCTTCA  
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCA  
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGCTATTGCTTGCCTTGGGACAGACCCTGTGGCTTAGGCT  
CTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGCTACAGTGGGCACAGCGCTGTGACGGCTCATGGGACT  
GTGCTGACGGCACAGATGAGGAGGACTGCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT  
CTGGTGCCACAGCCTGCTACCTGCCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA  
GACGCTGTGCGCATTTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCCTGTATGAGACGTGGGTGTGCC  
ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCATTACAG  
CTGCAGTCATTGGCAGCCTAGTGTGCGGCCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA  
TTTCGACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC  
CTTCTACGGGCAGCTCATTGCCCAGGGTGCCATCCCACCTGTAGAAGACTTTCTTACAGAGAATCCTAATGATA  
ACTCAGTGTGCGCAACCTGCGTTCTCTGTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCCAGGTG  
CCCGCCGTGCTCAGCGGGGCGCTTGATGCGACGCTGGTACGCCGTCTCCGCCGTGGGGCTTGCTCCCTCGAA  
CCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCCAGGTACACCTTCTGCTGCTCCCTTGAGGCCCTAGATG  
GTGGCACAGGTCCAGCCGTGAGGGCGGGGCAAGATGGGGAGCAGGCACCCCACTGCCCATCA  
AGGCTCCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC  
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCTGTTGCCCAGCCTGGGGCCCC  
CAGGACCAACCCGAGCCCCCTGGACCCACACAGCAGTCTGGCCCTGGAAGATGAGGACGATGTGCTACTGG  
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG  
CTCTACTGAGGCCCTCTCCCCGGGGGCTCTACTCATAGTGGCACAACCTTTTAGAGGTGGGTGAGCCTCCCCCTC  
ACCACTTCCCTCCCTGTCCCTGGATTTAGGGGACTTGGTGGGCCCTCCCGTTGACCCTATGTAGCTGCTATAAAGT  
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCCACTCCCT  
TCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTTTAAAAAGTAAAGTTCTTAGAGGATCATA  
GGTCTGGACACTCCATCCTTGCCAAACCTTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAAGTAGAGA  
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG  
CTCACAAAAAGAGTGCACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA  
GGAATCATACATCTC

## **FIGURE 68**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGT LQRPLVRDSRTSPANCTWLIL  
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP  
MGQGFLLSYSQDWLMCLQE EFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP  
VPSLPCNVTLEDFYGVFSSPGYTHLASVSH PQSCHWLLDPHDGRRLAVRFTALDLGF GDAVH  
VYDGP GPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHV RGYCLP  
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT  
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPD CADGSDEWDCS  
YVLPRKVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY  
GQLIAQGAIP PVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGR LMRRLVR  
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTG PAREGGAVGGQDGEQAPPLPIKA  
PLPSASTSPAPTTVPEAPG PLPSLPLEPSLLSGVVQALRGRL LPSLGPPGPTRSPPGPHTAV  
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

**Important features:**

**Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 442-462

**LDL-receptor class A (LDLRA) domain proteins**

amino acids 411-431, 152-171, 331-350 and 374-393

## **FIGURE 69**

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG  
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA  
GACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGG  
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC  
AAGCCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATA  
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGTTATTTTGGCCTTTGCTTGATAT  
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG  
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGACAGCAGTATGCTGTCTT  
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA  
GCCTGTGCATGAAAAAAAAAGAAGTTTTGTAATTTTTATATTACTTTTTTAGTTTGATACTAAGT  
ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAAAAA

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG  
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA  
GACGCGATGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGG  
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC  
AAGCCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATA  
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGTTATTTTGGCCTTTGCTTGATAT  
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG  
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGACAGCAGTATGCTGTCTT  
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA  
GCCTGTGCATGAAAAAAAAAGAAGTTTTGTAATTTTTATATTACTTTTTTAGTTTGATACTAAGT  
ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 70**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL  
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD  
GALIYRKLLFNPSGPYQKKPVHEKKEVL

### **Important features:**

#### **Potential type II transmembrane domain:**

amino acids 26-42

#### **Other potential transmembrane domain:**

amino acids 44-65, 81-101 and 109-129

#### **Leucine zipper pattern**

amino acids 78-99 and 85-106

#### **N-myristoylation site.**

amino acids 110-115

#### **Ribonucleotide reductase large subunit protein**

amino acids 116-127

[illegible]

GGGCGAGAAGTAGGGGAGGGCGTGTTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACCT  
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA  
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG  
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAGCCCC  
TGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTTCATACTTTTAT  
ATGTACTCAGACTTGATCGATTAATGAAGTGGTATTTTTGGCCTTTGCTTGATATTATCAAC  
TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCA  
AACATTGACAGTTGGTGGAGGGGTGTTTGCACCTTGTGACAGCAGTATGCTGTNTTGCCGAC

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

CAGCCCCGCGCGCCGGCCGAGTCTGCTGAGCCGCGGCTGCCGACGGGACGGGACTAGGCTGGGCGCGCCCCCGGGCCCCGCCGTGGGCATGGGGCGCACTGGCCCCGGGCGCTGCTGCTGCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGGACCCGGGACCCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGCGCGCCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGGCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATAACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAAATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGAAGCTGTGGCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAACTCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATAGCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCGCTTTCTCCTGTGCCACCCGTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTCACGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAAAAAAAAAACTTCATTCTAA

## **FIGURE 73**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL  
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS  
YIDTYFDTERSSTYRSKGFDTVTKYTQGSWTGFGEDLVTIPKGFNTSFLVNIATIFESENF  
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS  
LVLGGIEPSLYKGDIIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLLR  
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSFR  
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA  
EIIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFRCL  
QRRPRDPEVVNDESSLVRHRWK

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domain:**

amino acids 466-494

#### **N-glycosylation sites.**

amino acids 170-173 and 366-369

#### **Leucine zipper pattern.**

amino acids 10-31 and 197-118

#### **Eukaryotic and viral aspartyl proteases**

amino acids 109-118, 252-261 and 298-310

## FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG  
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCTGGGCGGGAGCCGGGAGGCGCGGCC  
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA  
CAACCTGGTGAAGGCCCCGCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG  
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG  
CGCGTGGTGTGGCCTGCCGCAGCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA  
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC  
GGGCCTTTGCCACTGCCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC  
GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAACCTGCTGCTTCGGGTGAACCATAT  
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCTAGCCGCG  
TGGTGGTGGTAGCCTCAGCTGCCCCTGTGCGGGACGTCTTGACTTCAAACGCCTGGACCGC  
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT  
GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC  
CAGGGCCTGTGAACTCGGAGCTGTTCTGCGCCATGTTCTGGATGGCTGCGCCCACTTTTG  
CGCCCATTTGGCTTGGCTGGTGTCTCCGGGCACCAAGAGGGGGTGCCAGACACCCCTGTATTG  
TGCTCTACAAGAGGGCATCGAGCCCCCTCAGTGGGAGATATTTTGCCAACCTGCCATGTGGAAG  
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG  
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC  
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCCTGAGGAGCCCACAGTTTCTCAACCTTACC  
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG  
CCTGAGATCCAGCTCTCCTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT  
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC  
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT  
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT  
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG  
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG  
TGAACCTGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCATGTTAATGAAGCG  
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG  
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA  
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCGGGGCGG  
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCCTCGGAATAAAGCGCGTTGACCGCCAAAA  
AAAAAAAAAAAAAAAAAAAA

1000 500 100 50 10 5

## **FIGURE 75**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSIGKMTALELARRGAR  
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG  
ISSCGRTREAFNLLLRVNHIGPFLLTHTLLLPCLKACAPSRVVVVASAAHCRGRLDKRLDRP  
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFRLRHVPGWLRPLLR  
PLAWLVLRAPRGGATPLYCALQEGIEPLSGRYFANCHVEEVPPAARDRAAHLWEASKRL  
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPOSSPDLSKMTHRIQAKVEP  
EIQLS

**Important features:**

**Signal peptide:**

amino acids 1-16

**Glycosaminoglycan attachment site.**

amino acids 46-49

**Short-chain alcohol dehydrogenase family**

amino acids 37-49 and 114-124



## **FIGURE 77**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHHLLP  
DGTLLLLQPPARGHAHDGQALSTD LGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM  
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAEKSDEGTYMCV  
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENTLLNPDPAEGPKPRPAVWLSWKV  
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG  
PDSNVLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP  
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGPW  
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD  
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST  
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA  
WKAKKKQELQHANS SPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS  
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS  
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY  
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA  
SLVSSSDGSFLADAHFARALAVAVDSFGFGLPREADCVFIDASSPPSPRDEIFLTPNLSLP  
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPVVDYS

### **Important features:**

#### **Transmembrane domain:**

amino acids 448-467

#### **N-glycosylation sites:**

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

#### **N-myristoylation sites.**

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

#### **Phosphotyrosine interaction domain proteins**

amino acids 740-753

DATE	TIME	LOCATION	WIND DIRECTION	WIND SPEED	TEMPERATURE	HUMIDITY	CLOUDS
1960-07-18	0800	STATION	090	10	28.0	75	BKN050
1960-07-18	1200	STATION	090	12	30.0	70	BKN050
1960-07-18	1600	STATION	090	15	32.0	65	BKN050
1960-07-18	2000	STATION	090	18	34.0	60	BKN050
1960-07-19	0400	STATION	090	10	28.0	75	BKN050
1960-07-19	0800	STATION	090	12	30.0	70	BKN050
1960-07-19	1200	STATION	090	15	32.0	65	BKN050
1960-07-19	1600	STATION	090	18	34.0	60	BKN050
1960-07-19	2000	STATION	090	20	36.0	55	BKN050
1960-07-20	0400	STATION	090	15	32.0	65	BKN050
1960-07-20	0800	STATION	090	18	34.0	60	BKN050
1960-07-20	1200	STATION	090	20	36.0	55	BKN050
1960-07-20	1600	STATION	090	22	38.0	50	BKN050
1960-07-20	2000	STATION	090	25	40.0	45	BKN050

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT  
CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCAAGGGGACACTGTGT  
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT  
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAAGAAGAAGGCCAGGAGACAAT  
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA  
ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG  
TCTTTACTGATCTCTCTGTTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCTTCTCCCAC  
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC  
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG  
GCTGAGGCCCCCTCCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC  
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC  
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG  
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCCTGGTGCTGCTGAGCCTTCTGTGAGC  
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA  
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC  
CCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA  
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT  
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCAGAAAGCTTTCACCTCAGCCTCAGAG  
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGCAATGTTCCA  
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCGAGAGCGGTGGCCTTGCTCTTCCGGCTG  
GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA  
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGC  
CTCATGCCCAGTGTGCGACCCTGCCTTCTCCCACTCCAGACCCCCACCTTGTCTTCCCTCCC  
TGGCGTCTCTCAGACTTAGTCCCAGGTTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT  
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT  
GTGAAAAACGTGATTCTTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG  
GACTCTGAATTCTAACATGCCCAGTGACTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG  
ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC  
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG  
TCCAGGCCTTGGTGCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG  
TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCCTTTGGAAAAAATGATGAAGA  
AAACCTTGGCTCCTTCCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA  
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG  
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA  
GCACAACCTACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT  
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAAGTGATT  
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT  
TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCCTGAC  
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG  
TCTGGCCCTATTTCCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAACTTGGAAG  
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT  
TATTTTCGTTTTGTGTACTTCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTCTT  
TTTACAGAGCAATTATCTTGTATATACAACCTTTGTATCCTGCCTTTTCCACCTTATCGTTCC  
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTATAAATAAAATGTTTCATCA  
GCTGCATAAAAAAAAAAAAAA

## **FIGURE 79**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPTEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS  
GTIYAE EEGQETMKGRVSIRD SRQELSLIVTLWNLTLQDAGEYWCGVEKRGPD E SLLISLFV  
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG  
TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI  
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE E KEAPSQAPEGD  
VISMPPLHTSEEELGFSKFVSA

### **Important features:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domain:**

amino acids 248-269

#### **N-glycosylation site.**

amino acids 96-99

#### **Fibrinogen beta and gamma chains C-terminal domain.**

amino acids 104-113

#### **Ig like V-type domain:**

amino acids 13-128

姓名	性别	年龄	籍贯	职业	文化程度	健康状况	婚姻状况	子女情况	其他
王德胜	男	45	山东	工人	小学	良好	已婚	2子1女	
李秀英	女	38	河北	农民	初中	良好	已婚	1子1女	
张国强	男	52	河南	干部	高中	良好	已婚	2子1女	
刘小红	女	28	江苏	教师	大学	良好	未婚	无子女	
陈为民	男	60	浙江	退休	小学	一般	已婚	3子2女	
赵大伟	男	35	湖北	商人	高中	良好	已婚	1子1女	
孙丽娟	女	42	湖南	护士	中专	良好	已婚	2子1女	
周建民	男	58	四川	工人	小学	一般	已婚	3子2女	
吴小芳	女	32	广东	文员	初中	良好	已婚	1子1女	
郑为民	男	48	山西	农民	小学	一般	已婚	2子1女	
冯大刚	男	55	陕西	干部	高中	良好	已婚	2子1女	
马秀珍	女	40	甘肃	教师	大学	良好	已婚	1子1女	
徐国强	男	30	福建	商人	初中	良好	已婚	1子1女	
林小红	女	25	江西	文员	高中	良好	未婚	无子女	
周为民	男	65	安徽	退休	小学	一般	已婚	3子2女	
吴大伟	男	38	广西	工人	初中	良好	已婚	1子1女	
孙丽娟	女	45	云南	护士	中专	良好	已婚	2子1女	
周建民	男	50	贵州	干部	高中	良好	已婚	2子1女	
吴小芳	女	35	海南	商人	初中	良好	已婚	1子1女	
郑为民	男	40	重庆	工人	小学	一般	已婚	2子1女	
冯大刚	男	55	四川	干部	高中	良好	已婚	2子1女	
马秀珍	女	40	湖南	教师	大学	良好	已婚	1子1女	
徐国强	男	30	湖北	商人	初中	良好	已婚	1子1女	
林小红	女	25	江西	文员	高中	良好	未婚	无子女	
周为民	男	65	安徽	退休	小学	一般	已婚	3子2女	
吴大伟	男	38	广西	工人	初中	良好	已婚	1子1女	
孙丽娟	女	45	云南	护士	中专	良好	已婚	2子1女	
周建民	男	50	贵州	干部	高中	良好	已婚	2子1女	
吴小芳	女	35	海南	商人	初中	良好	已婚	1子1女	
郑为民	男	40	重庆	工人	小学	一般	已婚	2子1女	
冯大刚	男	55	四川	干部	高中	良好	已婚	2子1女	
马秀珍	女	40	湖南	教师	大学	良好	已婚	1子1女	
徐国强	男	30	湖北	商人	初中	良好	已婚	1子1女	
林小红	女	25	江西	文员	高中	良好	未婚	无子女	
周为民	男	65	安徽	退休	小学	一般	已婚	3子2女	
吴大伟	男	38	广西	工人	初中	良好	已婚	1子1女	
孙丽娟	女	45	云南	护士	中专	良好	已婚	2子1女	
周建民	男	50	贵州	干部	高中	良好	已婚	2子1女	
吴小芳	女	35	海南	商人	初中	良好	已婚	1子1女	
郑为民	男	40	重庆	工人	小学	一般	已婚	2子1女	
冯大刚	男	55	四川	干部	高中	良好	已婚	2子1女	
马秀珍	女	40	湖南	教师	大学	良好	已婚	1子1女	
徐国强	男	30	湖北	商人	初中	良好	已婚	1子1女	
林小红	女	25	江西	文员	高中	良好	未婚	无子女	
周为民	男	65	安徽	退休	小学	一般	已婚	3子2女	
吴大伟	男	38	广西	工人	初中	良好	已婚	1子1女	
孙丽娟	女	45	云南	护士	中专	良好	已婚	2子1女	
周建民	男	50	贵州	干部	高中	良好	已婚	2子1女	
吴小芳	女	35	海南	商人	初中	良好	已婚	1子1女	
郑为民	男	40	重庆	工人	小学	一般	已婚		

TTGTGACTAAAAGCTGGCCCTAGCCAGGCCAGGGAGTGACGCTGCAGGCGTGGGGGTGGCAGGA  
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC  
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT  
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA  
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAA  
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG  
CAGGGCTCTCAGAAGGCGGTGGTGCCAGCTGGGATCATGTGTGTTGGCCCTGGTCTGTCTGC  
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG  
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC  
TTATTTTACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA  
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC  
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC  
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCC  
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCA  
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGAAGACAAGCCAGCGAATA  
AAGGATGGTTGAACGTGAAA

[illegible]

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187
```

&lt;subunit 1 of 1, 146 aa, 1 stop

&lt;MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD  
YEADGSTNNGIFQINSRRWCSNLTNPVNPVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGY  
WEAWRHHCQGKDLTEWVDGCDP

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homolgous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

## FIGURE 82

AGCCGCTGCCCCGGGCGGGCGCCCCGCGCGGCACCAATGAGTCCCCGCTCGTGCCTGCGTTC  
GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA  
AGCTGTCGTCGGTGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC  
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC  
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAAGTGTCCACAC  
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTG  
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGCAGCAGTGGGGAGCT  
GGAGAAGTGCGGCTGTGACAGGACAGTGCATGGGGTCAGCCACAGGGCTTCAGTGGTCAG  
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG  
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG  
GAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCACGGGGTGTGAGGCTCCTGTG  
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG  
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC  
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGAGCCTAGCCCCG  
ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGTGGGCACGAGGGGCCGCACATGCAACAAG  
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA  
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCACTGGTGTGCTTCGTCAAGTGCCGGC  
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATTGACCGCCTGCCTAGCCCTGCGCCGGC  
AACCACCTAGTGGCCCAGGGAAGGCCGATAATTTAAACAGTCTCCCACCACCTACCCCAAGA  
GATACTGGTTGTATTTTTTGTCTGGTTTGGTTTGGGTCCTCATGTTATTTATTGCCGAA  
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT  
GCCACTGACCAAGGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA  
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT  
TACCACCACATGGCTACTGACCGTGTATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA  
GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAGAACTCTTAACTCTCCAGCACACA  
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG  
GAACAAGCAGATACCAGGTCAAGGGCACCAAGGTTTCAATTCAGCCCTTACATGGACAGCTAGA  
GGTTCGATATCTGTGGGTCCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA  
GTCCCACCCTAGAACCCAGCCTGCCCCAGCCTGCCCCCTGGGAAGAGGAACTTAACCACTCC  
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT  
TGCAGTCATGCCCCAGTCACCTTTCACAGCGCTGTTCTCCATGAACTGAAAAACACACAC  
ACCTGCGAGA  
GAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCCCAGTCACCTTTCACAGCACTGTTCTC

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

&lt;subunit 1 of 1, 351 aa, 1 stop

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE  
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLPPVFGKVVTQGTREAAFVYAISSAGVAFV  
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSNIAYGVAFSQSFDVVRERSKGASSSRALM  
NLHNNEAGRKAILTMRVECKCHGVSGSEVKTCTWRVPPFRQVGHALKEKFDGATEVEPRR  
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC  
CGRGFHTAQVELAERCSCKFWCCFVKCRQCRLVELHTCR

Signal peptide:

N-glycosylation sites.

Wnt-1 family signature.

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGCTGCCTGCAT  
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC  
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGCCCTGGCTGTCCTGGTC  
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC  
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGGCGCTGG  
GTGCCCTGAAGGAGGAGGTTCGAGAGCTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG  
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT  
GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG  
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGGCCGTGAGGCTCCAGAACAACCTCCTGCGAGCCG  
TGCCCCACGTGCTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC  
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC  
TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG  
GCTGTGCGCCATCTGGGCAAGGTTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT  
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTCATGATGC  
TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG  
AAAAGGCACAACCTGCTTGACCCCGCCCAGTGCCCTGGAGCCGCGCCCATTCAGCATGTCGTA  
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCT  
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC  
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC  
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA  
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGA  
AGCTGTTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

[illegible]

&lt;subunit 1 of 1, 293 aa, 1 stop

MDTTRYSKWGGSSSEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA

ELRERVTOGLAEAGRGREDVRTELFRALAVRLQNNSCPCPTSWLSFEGSCYFFSVPKTTW

HWNQGEPNDAWGRENVMMLHTGLWNDAPCDSEKDGWICEKRHNC

## Type II transmembrane domain:

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287



Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

><MW: 35844, pI: 5.45, NX(S/T): 2

Important features:

amino acids 1-26

## FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGAT  
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGACTCGTAGGCGGA  
CAGCTTTAGTGCCGCGCCGCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC  
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAATA  
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT  
ATCCAACCTTTGTTTGGAAAGCTTATTATGACAATACCATTTTTTCATAGAGTTGTGCCTGGTTT  
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC  
CATTCAAAGATGAATTTCAATTCACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA  
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTCGAGCAGATGAACT  
TAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC  
TGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT  
GAGGTTTTGTTTAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAGA  
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT  
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG  
GGCAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT  
TGTAAGAAAGTAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG  
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA  
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC  
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACTCTTAGCAG  
CAAAACAAAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA  
GCCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG  
GAAGCAACAGTCAAAGAAGGGAACCTTCCCGGGAAGATCAGACCTTGCAGTCTGAACCAGT  
TTAAATCTAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA  
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT  
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA  
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATATAAAT  
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG  
TTCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG  
AAAAACAATTATCTTGTTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGTTTACTGG  
TACATGTGTTTTTTCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCCT  
TCCACAAAAA

[illegible]

```
><subunit 1 of 1, 472 aa, 1 stop
```

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYDNTIFHRVVPGFI  
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFFTLGRADELN  
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEK  
PEEEVKKLKPKGTKNFSLLSFGEAAEEEEEEVN RVSQSMKGKSKSSHDLKDDPHLSSVPVV  
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKDTSANVKSAGEGEVEKKSV  
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAAPPDGAVA EYRREKQKYEALRK  
QQSKKGTSREDQTLALLNQFKSKLTQAI AETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK  
DASMQSDTFEIIYDPRNPVNKR RREESKKLMREKKERR

Signal peptide:

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

## FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCTGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG  
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTCGCCTTGCTGACGGCGTCG  
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC  
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA  
CCCTTCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG  
CTCCTTCAAGTGGTTTTGGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA  
GGAGGAACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC  
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCT  
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGGAGCCCTGGA  
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTGCGCAGGCTGGAG  
TGCAGTGGCAGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC  
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTGGGGCCAGCTGGCCCTCG  
ATGTACGTCAGCACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT  
GAACGTGTACGCGCCGGCGCGCGCCCGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC  
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC  
GAGAAAGTGGTGCTGGTGTTCCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGA  
CGACAGCCACGCGCGCGGGAACCTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGC  
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCCTGTTTCGGCCAGTCGGCG  
GGGGCCATGAGCATCTCAGGACTGATGATGTCACCCCTAGCCTCGGGTCTCTTCCATCGGGC  
CATTTCCCAGAGTGGCACC GCGTTATTTCAGACTTTTCATCACTAGTAACCCACTGAAAGTGG  
CCAAGAAGGTTGCCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC  
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTCCA  
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG  
TGATCCCAGATGACCCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT  
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCCCTTATAATATCACCAAGGAGCAGGT  
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA  
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC  
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA  
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC  
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC  
TATGCAGGAAGGAGCCAAAGAGGGGTTTGGCCCCACCATCCAGGCCCTGGGGAGACTAGCCA  
TGGACATACCTGGGGACAAGAGTTCTACCCACCCAGTTTGAAGTGCAGGAGCTCCCTGCT  
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC  
TGACATCCCATGATGCCCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC  
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTCCCTTCTTCAAATCCT  
CCCACCCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTGACCCAGACTGCCACTGC  
CCCTGTCACTGCACCCAGCTTGGCATTACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT  
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACCTTTTGGTAGTTTGGGA  
TCTTCTCTCCACCCACACTTATCTCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG  
TCTCTTCAATAAAGAAGTGTTGATTAGAAAAA

594556 073001

## FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSSAPSSG  
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVFPFSRPP  
LGILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS  
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVEYAPARAPGDPQLPVMVWFPGGAF  
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA  
AFGGDPGNVTFLGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA  
HLACNHNSTQILVNCLRALSGTKVMRVSNMKMRFLQLNFQRDPEEIIWSMSPVVDGVVIPDD  
PLVLLTQQKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMMD  
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

### **Important features:**

#### **Signal peptide:**

amino acids 1-29

#### **Carboxylesterases type-B serine active site.**

amino acids 312-327

#### **Carboxylesterases type-B signature 2.**

amino acids 218-228

#### **N-glycosylation sites.**

amino acids 318-321, 380-383 and 465-468

4002705957550

[illegible][illegible]

## FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSPYPRQDWTGSTPAYGYWFK  
AVTETTTKGAPVATNHQSREVEEMSTRGRFQLTGDPKGNCSLVIRDAQMQDESQYFFRVERGS  
YVTYNFMNDGFFLKVTVLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS  
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL  
GLELPGVKAGDSGRYTCAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL  
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQSPDPGVLELPRVQVEHEGEFTCHARHP  
LGSQHVLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR  
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP  
KSSTQAPESQESQEELHYATLNFPFGRPRPEARMPKGTQADYAEVKFQ

### Important features:

#### Signal peptide:

amino acids 1-15

#### Transmembrane domain:

amino acids 399-418

#### N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

#### Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

# FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAAGTGTATTTACTGCTGCGTTT  
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACCTCTCAAACAAAGA  
AAGTCAAGCAGCCAGTGCATCTCATTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT  
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA  
CAATGGAAACAATTCCTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA  
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC  
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA  
GTTTGTCTCATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG  
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT  
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC  
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC  
TGCAAGATGAGTATTGGGTAATCATTCAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG  
TCTGGAACAACAAGTGTATTAATTAACCTTTCAGATGTTAATGACAATAAGCCTATATTTAA  
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA  
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT  
GATTCGCAAAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTTAA  
AAAGAAAGTGGATTTTGGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC  
ATGTTCTCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTCATTAAGATCCAG  
GTGGAAGATGTTGATGAGCCTCCTCTTTTCTCCTCCATATTATGTATTTGAAGTTTTTGA  
AGAAACCCACAGGGATCATTGTTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT  
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCATATCAATGATAATGGTACAATC  
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC  
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCCTGTATGTGCAAGTTCTTAACATCA  
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAGGCTCT  
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATT  
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAAGTTTACAATCATAGATAATCAAG  
ATAACACAGCTGTCAATTTGACTAATAGAACTGGTTTTAACCTTCAAGAAGAACCTGTCTTC  
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCCTTAC  
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG  
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTCATTATGATCATA  
TTTGGGTTTATTTTTTTGACTTTGGGTTTAAACAACGAGAGAAAACAGATTCTATTTCTGA  
GAAAAGTGAAGATTTTCAGAGAGAATATATTTCCAATATGATGATGAAGGGGTGGAGAAGAAG  
ATACAGAGGCCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT  
CGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA  
CAGTGCCATATTCAGGAAATTCATTCTGGAAGAGCTCGAAGAAGCTAATACTGATCCGTGTG  
CCCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC  
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT  
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTCTCTGCAGTGCAGTCAATAATTAGG  
GCTTTTTTACCATCAAAATTTTTTAAAGTGCTAATGTGTATTTCGAACCCAATGGTAGTCTTAA  
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC  
CTGGAGTAAATACTCCATGGTTATTTTTAAGCTACCTACATGCTGTCAATTGAACAGAGATGTG  
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAATAATG  
TAGGAAGATATTTAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT  
CATTATTTACTTAGGAAAGAGTAAAAATACCAAACGAGAAAAATTTAAAGGAGCAAAAATTTG  
CAAGTCAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA  
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTGTTTCCTTTAT  
TTAAA

094655 073004

## FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLPCLGATENSQTKKVKQPVRSHLRVKGWVWNQFFVPEEMNTTSHH  
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT  
GRAVEPESEFVIKVS DINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL  
LYSLLQGQPYFSVEPTTG VIRISSKMDRELQDEYVWIIQAKDMIGQPGALSGTTSVLIKLS D  
VNDNKPIFKESLYRLTVSESAPTGT SIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIIITNHE  
TQEGIVILKKKVDFEHQNHYGIRAKVKNNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL  
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA  
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR  
DESIEEHFHYFNLSVEDTNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP  
SLTSTNTLTIHVCDGSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ  
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKT TSAEIRSLY  
RQSLQVGPD SAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGLAGSLSSLES AVSDQD  
ESYDYLNELGPRFKRLACMFGSAVQSN

### Important features:

#### Signal peptide:

amino acids 1-21

#### Transmembrane domain:

amino acids 597-617

#### N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,  
516-519 and 534-537

#### Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

100229-53906

**FIGURE 96**

ATTTCAAGGCCAGCCATATTTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTTNTA  
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT  
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAACTTTCAGATGTTAATGACAA  
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA  
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC  
AGCATTGAAGAGGATGATTGCGAAACATTTGACATTATT

[illegible]

## FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCAACCCCGAC  
CCAGAGCTTCTCCAGCGGCGGCGCAGCGAGCAGGGCTCCCCGCCTTAACCTCCTCCGCGGGG  
CCCAGCCACCTTCGGGAGTCCGGGTTGCCACCTGCAAACCTCTCCGCCTTCTGCACCTGCCA  
CCCCTGAGCCAGCGCGGGGCCCCGAGCGAGTCAATGGCCAACGCGGGGCTGCAGCTGTTGGGC  
TTCATTCTCGCCTTCTCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG  
GATTTACTCCTATGCCGCGGACAACATCGTGACCGCCCAGGCCATGTACGAGGGGCTGTGGA  
TGTCTGCGTGTGCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT  
CTGAGCAGCACATTGCAAGCAACCCGTGCCCTTGATGGTGGTGGCATCCTCCTGGGAGTGAT  
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGC  
AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTCTTGCAAGGTCTGGCTATTTTA  
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCCCAGT  
CAATGCCAGGTACGAATTTGGTCAAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCGC  
TTCTGGGAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCAACACCA  
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTGTGACACAGAGGCAAAAG  
GAGAAAATCATGTTGAAACAAACCGAAAAATGGACATTGAGATACTATCATTAACATTAGGAC  
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAAA  
ACCCATGTGTTAAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCTCA  
ATATAGGAGGGAAGATTTTTCCATTTGTATTACTGCTTCCCATTTGAGTAATCATACTCAAAT  
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA  
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAAATATCTCTAAAT  
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTCGTCC  
TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGCTTTGGGTGCCTTTG  
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT  
CATATACTTATTTTATTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT  
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTA  
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTTCTGCATG  
ACCAAAGTGATAAATTCCTGTTGACCTTCCACACAATCCCTGTACTCTGACCCATAGCACT  
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCAGGTGTTGT  
AACACAACCTTATTGATTGAATTTTAAAGCTACTTATTCATAGTTTATATCCCCCTAACT  
ACCTTTTTGTTCCTTCCCTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTA  
TATCTTCCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA  
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT  
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA  
TTAGTTTATATTACTCTTATTCTTTGAACATGAACCTATGCCTATGTAGTGTCTTTATTTGCT  
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT  
CACTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT  
GTGGTTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC  
CTTCATGTGGCTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTCT  
GACATGTTTGTGCTCTGTTCCATTTTAACTGCTCTTACTTTTCCAGTCTGTACAGAATG  
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGTGTTGGCACTGGTGTCTGGAGACCTG  
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTAA  
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCCACCTCACAGTGATG  
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAAGTGCTAT  
ACTAAGGGAAAGAATTGAGGAATTAAGTGCATACGTTTTTGGTGTGCTTTTCAAATGTTTGA  
AAATAAAAAAATGTTAAG

091000 073004  
000520 000000

## **FIGURE 98**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI LAFLGWIGAI VSTALPQWRIYSYAGDNIVTAQAMY EGLWMSCVSQSTGQI  
QCKVFDSLNLSS TLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDDEVQKMRMAVIGGA  
IFLLAGLAILVATAWYGNRIVQEFYDPMTPVNARYEFGQALFTGWAAASLCLLGALLCCSC  
PRKTTSYPTPRPYPKPAPSSGKD YV

### **Important features:**

#### **Signal peptide:**

amino acids 1-21

#### **Transmembrane domains:**

amino acids 82-102, 118-142 and 161-187

#### **N-glycosylation site.**

amino acids 72-75

#### **PMP-22 / EMP / MP20 family proteins**

amino acids 70-111

#### **ABC-2 type transport system integral membrane protein**

amino acids 119-133

[illegible]

1. The first step in the process of developing a new product is to identify a market need. This involves conducting market research to determine what consumers want and need. Once a need is identified, the next step is to develop a concept that addresses this need. This concept should be unique and offer a clear benefit to the consumer.

## FIGURE 100

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG  
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG  
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT  
GGATGTCCTGCGTGTTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT  
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG  
TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA  
GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTA  
TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

100-20-533160

[illegible]

GGGCCCGACCATTTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC  
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC  
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANACCGGGCAGATCCAGTGCAA  
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT  
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTTGGCATGAAGTGTATGAAGT  
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT  
CTTGCAAGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT  
TCTATGACCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC  
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGCGA

## FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC  
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT  
GCNTGTGCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG  
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC  
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG  
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC  
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG  
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG  
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

05  
14  
20  
24  
27  
30  
33  
36  
39  
42  
45  
48  
51  
54  
57  
60  
63  
66  
69  
72  
75  
78  
81  
84  
87  
90  
93  
96  
99  
102

## FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCCTTGCTGAATCTGAGCAGCACATTNC  
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG  
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT  
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGACAGGTCTGGCTATTTTAGTNCCACAGCAT  
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCAGTCAATGCCAGGTAC  
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC  
CCTACTTTGCTGTTCTGTCCCGAA

10029-43940

**609**

*[The page contains extremely faint, illegible horizontal lines, likely representing a blank or heavily degraded document.]*

## FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA  
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCAGTCAATGCCAGGTACGAAT  
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA  
CTTTGCTGTTCTG

10020 5530 460

**00000000**

TTCTTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC  
TGGCGAACAAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTG  
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT  
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG  
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT  
GGCTGTCAATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGNNGCCACAGCAT  
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCCAGTCAATGCCAGGTA  
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG  
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

**FIGURE 107**

Figure 1 illustrates the stages of chick development from fertilization to hatching. The diagrams show the progression from a single cell to a fully formed chick, including the formation of the embryo, the yolk sac, and the hatching process.

## FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCTGCGCCTGCAC  
 CGCGTAGACCGACCCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCCGCCCCTGCCCCG  
 ACCGGTCCCCGCCTTTTGTAAAACTTAAAGCGGGCGCAGCATTAAACGCTTCCCCGCCCCGGT  
 GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAACATGGCGAAGGTGGAG  
 CAGGTCCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCGAGGTCCCTTCACCGATGTTGT  
 CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTAAAGGTGAAGACTA  
 CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCCTCAATT  
 AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT  
 TATGGTTCAGTCTATGTTTGCTCCAACCTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG  
 CAAAACCGGAAGACCTTATGGATTCAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT  
 GATAAACCATATGATGTAGAAATAAATAAAATTATATCCACAACTGCATCAAAGACAGAAAC  
 ACCAATAGTGTCTAAGTCTCTGAGTTCCTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG  
 AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAAGCAGTTCAAG  
 GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTTCAGCATTAGC  
 CCAAACCTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGCTCTGGTGTTTGTCTTTA  
 TCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAGAGGTAGCATGCACAGGATGGTAAATTG  
 GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT  
 AATGTATGATGACATCTCACAGGTCTTGCTTTAAATTACCCCTCCCTGCACACACATACAC  
 AGATACACACACAAAAATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG  
 ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT  
 GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC  
 TCTTAAATGACACCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT  
 GCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCCAGTGGCCCACTCCCGGCCAGGCTG  
 CTTTCCGTGTCTTCAGTTCGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA  
 AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTAATCGTCATAAGTGAGAGGCGTGTGT  
 TGAATGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA  
 GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAATGTTATTTCAGAGATGTTTAAATGCATA  
 TTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTGACAAGAGTACAGTTAATGC  
 TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT  
 GTCTCTGGAGAGTCTGGTTCATGTGGAGGTGGGGTTTATTGGGATGCTGGAGAAGAGCTGCCA  
 GGAAGTGTTTTTCTGGGTGAGTAAATAACAACCTGTCATAGGGAGGGAAATTCTCAGTAGTG  
 ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA  
 CCACCTCTCAACCATTACTCACACTTCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC  
 TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC  
 ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTAC  
 TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA  
 CTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

090405 07304

## **FIGURE 109**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID  
AGASINVSVMQLQPFDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRVFE  
LPAENDKPHDVEINKIIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE  
NKQFKEEDGLRMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

### **Important features:**

#### **Transmembrane domain:**

amino acids 224-239

#### **N-glycosylation site.**

amino acids 68-71

#### **N-myristoylation site.**

amino acids 59-64, 64-69 and 235-240

[illegible]

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTC AACCANTACTCACATTTTCNAGCGCCAG  
 GTCCANGTCTGAGCCTGACTTCCCCTTG GGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG  
 CTGCAGAGGNTTAGAAGCGAGGGCAC CAGCAGTTTTGGGTGGGGAGCAAGGNGAGAGAAA  
 CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA  
 AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC  
 CAAAATTAAGAATTCTTTTGTCATTTTGTCACATTTGCTCTATGGGGGGAATTATTATTTT  
 ATCATTTTTTATTATTTTGCCATTGGAAGGTTAACTTTAAATGAGC

## FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT  
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG  
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT  
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC  
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA  
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT  
AAATTGTATTGGTTCATGTAGTGAAGTCAAAGTGTATTTCAGAGATGTTTAATGCATATTTA  
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG  
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

00223 433469

**FIGURE 112**

CCCTGGTGGTTTTGTTCCTTAATTCTGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA  
GNATGCACCNCGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT  
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT  
ACCCATCCCTGNANACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG  
AAAGTTAAAAATGTATAGTAAC

[illegible]

# SECRET

1. The first part of the document is a title page. It contains the title of the document, the author's name, and the date of the document.

2. The second part of the document is an introduction. It provides a brief overview of the document's content and the author's purpose in writing it.

3. The third part of the document is the main body. It contains the main content of the document, which is organized into several sections.

4. The fourth part of the document is a conclusion. It summarizes the main points of the document and provides a final statement.

5. The fifth part of the document is a bibliography. It lists the sources of information used in the document.

6. The sixth part of the document is an appendix. It contains additional information that is related to the main content of the document.

7. The seventh part of the document is a glossary. It defines the terms used in the document.

8. The eighth part of the document is an index. It provides a list of the topics covered in the document and the pages where they can be found.

9. The ninth part of the document is a list of figures. It provides a list of the figures included in the document and the pages where they can be found.

10. The tenth part of the document is a list of tables. It provides a list of the tables included in the document and the pages where they can be found.

## FIGURE 114

TGCTTTC CGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC  
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG  
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC  
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAAGTGTATTTCAGAGATGTTTAAATGC  
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA  
TGCTGCGTGC

40529.200

[illegible]

AAACCTTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACNTGNGT  
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC  
TGGATTACCTCTCTTAAAATGACACCCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN  
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANCCCC  
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGTCTCCTTGGGANTGATGA  
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA  
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT  
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG  
TTTAATGCATATTTAANTTATTTAATGTATTTCATNTCATGTTTTCTTATTGTCACAAGGGT  
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

# THE BIBLE

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC  
CACGTGGCCCACTCCCGGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC  
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCCAAAGGAATTGCACTGTGGCAGCATCAGACG  
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC  
AGTGCTTTGTTCACCTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA  
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTTC  
TTATTGTCAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT  
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

**A** **B** **C** **D** **E**

CGAGACTCCGGGTGCTGTGGCCCGGCCTTGCGGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA  
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC  
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCA  
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTC AAGCGATCCTGCTGAGTAGCTGGGA  
CTACAGGACAAAATTAGAAGATCAAAATGGAAAAATATGCTGCTTTGGTTGATATTTTTTACC  
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA  
GGTACCCCGGATTGTCAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG  
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGG  
CTTTCTGAATTGGAGGATTATCTTTCCCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT  
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG  
TATCTGTTAGGAGAAAAGAGACAGGTGTATGGCACCGACAGCAGGTT CAGCATCTTGGACAAA  
AGGTTCTTAACCAATTTCCCTTT CAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT  
TCTCATTTCCCTCAGCATGTTCTAACTGCTGCCCACTGTGTT CATGATGGAAGGACTATG  
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG  
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGGTAC  
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAAATCTGGCCGGGGTC  
AGAGGATTGCCGAAGGGAGGCCCTTCTTT CAGTGGACCCGGGTCAAGAATACCCACATTCCG  
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGA CTATGACTATGCTCTTCTGGA  
GCTGAAGCGTGCTC AAAAAAGAAATACATGGAAC TTGGAATCAGCCCAACGATCAAGAAAA  
TGCCTGGTGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT  
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC  
GGGCTCCACCGGTT CGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC  
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTT CAGAAGGAC  
TACAACGTTGCTGTT CGCATCACTCCCC TAAAAATACGCCCAGATTTGCCCTCTGGATT CACGG  
GAACGATGCCAATTGTGCTTACGGCT TAA CAGAGACCTGAAACAGGGCGGTGTATCATCTAAA  
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCCTGGACTT  
GAACTCTGTCAATAGCATTTC AACATTTTTTCAAAATCAGGAGATTTTCGTC CATT TAAAAA  
TGTATAGGTGCAGATATTGAAACTAGGTGGGCAC TTCAATGCCAAGTATATACTCTTCTTTA  
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTTGCC TTCTTAAAAATTAGACACACTTT  
AAACCTTCAAACAGGTATTATATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC  
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA  
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA  
CTCTGAGATGGATCCATT CAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT  
GGGACATTTAGTTT TAGTTTTTTTGAAGAAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA  
CAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA  
TGGGAGAAATAGTTTGTTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT  
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGATGCATAGGG  
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATATTTTATGCTATACAT  
TCTATGTATGAGGTGCTACATTTTTTAGGACAAAGAACTCTGTAATCTTTTTTCAAGAAAGAGT  
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG  
TAATTTT TAGATATGTCCTTTCC TAAAAATGAATAAAATTTATGAATATGA

## **FIGURE 118**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLLIDGSEMEWDFMWHLLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC  
GIECQKELPTPSLSELEDYLSYETVVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV  
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAAHCVHDGKDYVKGSKKLRV  
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGRRRKKSGRGQRIAEGRPS  
FQWTRVKNTHIPKGWARGMGMDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS  
GFDNDRADQLVYRFCVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG  
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

### **Important features:**

#### **Signal peptide:**

amino acids 1-16

#### **N-glycosylation sites.**

amino acids 90-93, 110-113 and 193-196

#### **Glycosaminoglycan attachment site.**

amino acids 236-239

#### **Serine proteases, trypsin family, histidine active site.**

amino acids 165-170

## FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT  
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTAGTGTCCGATTCTGATTCCGGCAAGG  
ATCCAAGCATGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC  
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG  
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT  
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC  
TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA  
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA  
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG  
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA  
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA  
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT  
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAAGGTCCTGATCACTT  
ATATCTGGAACCAAAACCCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA  
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCAGAAATTTCCAGACAAAGAGATACTGAGA  
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTGTAACCTCGGGCTCCGCTGACAG  
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC  
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG  
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC  
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATAACAAGCAGATCATGC  
CTTATGACCTCTACCATCCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC  
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA  
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTACACCCTAAGATGCCCATCGCGCAGCCCT  
GCAACATTTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT  
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG  
CTGTAGCCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA  
AACCCAAAGAGAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA  
GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTAAGTTGTAAAAGCACAGACTGTTCTATA  
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA  
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA  
AAAAAAAA

## **FIGURE 120**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRCLS  
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPSLKCQ  
AKGTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR  
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDLHLYLETKTLQGTKGENSLSSTGTFL  
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS  
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD  
LYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI  
FDCPKWLAQEWSPCTVTGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK  
EKLPVEAKLPWFKQAQEEGA AVSEEPS

### **Important features:**

#### **Signal peptide:**

amino acids 1-25

#### **N-glycosylation site.**

amino acids 251-254

#### **Thrombospondin 1**

amino acids 385-399

#### **von Willebrand factor type C domain proteins**

amino acids 385-399, 445-459 and 42-56

## FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC  
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG  
GCCCCCGGTTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCCGGCCTCAT  
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCGCGCCACGGCCGCCG  
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG  
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT  
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG  
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG  
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA  
TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCTCAAAAGTTCAGCTCCCAGCA  
GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC  
AGTGAACAAAGCTATAATAAAAGCTTTTGTTATAGCCGGAGCAAAGTGGCTAACATTCTTTT  
TACCAGGGAAGTACCCCGCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG  
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC  
AATTTGGTGTGATGGGCTTTTTTCAAACTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT  
GGCCTCTTCACCTGAGGTAGAAGGAGTGTCAGGAAGATACTTTGGGGATTGTAAAGAGGAAG  
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAAGTCTGGGATATCAGTGAAGTG  
ATGGTTGGCCTGCTAAAATAGGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAG  
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG  
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT  
TGGGATAAGAGAATTTTCAAGAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA  
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA  
ATATTTGTCAGAATTAAGTGAAGTCAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT  
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC  
CTGGTGTGTGCACACAAGTCTTACTTGAATAAATTTACTGGTAC

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC

## **FIGURE 122**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE  
LLRLGARVIMGCRDRARAEAAAGQLRREL RQAAECGPEPGVSGVGELIVRELDLASLRSVRA  
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR  
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSLANILFTRELARRLEGTNVTNVNLHPG  
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDKKEEE  
LLPKAMDESVARKLWDISEVMVGLLK

### **Important features:**

#### **Signal peptide:**

amino acids 1-21

#### **Short-chain alcohol dehydrogenase family protein**

amino acids 134-144, 44-56 and 239-248

#### **N-glycosylation site.**

amino acids 212-215 and 239-242

## FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN  
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT  
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT  
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG  
TTACTGAAAAATTATTTTTGGGATAAGAGAATTCAGCAAAGATGTTTTAAATATATATAGT  
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATACTGGGCA  
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGAAGTCAAGTGCTATCGAGAGGTTT  
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT  
TGTGTGGAAATTATCTGCCTGGCTT

40640453450

**B**

GAGAGGACGAGGTGCCCGTGCCTGGAGAATCCTCCGCTGCCGTCCGGCTCCCGGAGCCCAGCC  
CTTTCCTAACCCAAACCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC  
CCCAGCGTTACCATGCATCCTGCCGTCTTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT  
GCTCCTGGTAACTTTGGGTTTTTACTCCTGTAACAACCTGAAATAACAAGTCTTGCTACAGAGA  
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT  
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT  
TCCAAATGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC  
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAG  
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTTGGCAGATTACATCAGGCAACAAAAAAG  
TGACCCCATTTCAAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAA  
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACCTATAGAGTTTTTTGAACGAGTAGCG  
AATATTTTTGCATGATGACTGTGCCTTCTTTCTGCATTTGGGGATGTTTCAAACCGGAAAG  
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT  
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTCT  
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT  
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC  
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT  
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCTCTGTAATCGCTATTGACAGCTTTAG  
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCTTGGA AAACTCAAGCAATTCCG  
TATTTGACTTACATTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT  
ACAGCCCCAGGAGAGCAAGCCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA  
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAAACTTG  
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA  
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA  
AAAAA

## FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ  
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMMKREYR  
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILH  
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNW IQDKCVPLVREI  
TFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH  
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG  
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

**Important features:**

**Signal peptide:**

amino acids 1-29

**Endoplasmic reticulum targeting sequence.**

amino acids 403-406

**Tyrosine kinase phosphorylation site.**

amino acids 203-211

**Thioredoxin family proteins**

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA  
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG  
GGATGATGATGAAGAGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTA

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

## **FIGURE 127**

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTGCGCCGCNKGAGCCCGGGTCGAGAGGACNAGG  
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGCGCTCCCGGAGCCCAGCCCTTTCCTAACCC  
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTNCGGANCCCAGCGTNACC  
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC  
TTGGGTTTTTACTCCTGTAACAACGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA  
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTGTTTTAGTCAG  
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA  
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA  
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG  
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

127

## FIGURE 128

GCCCACGCGTCCGATGGCGTTCACGTTGCGGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA  
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTGATGAGCTGAAGACTGAT  
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTAAGTCCCAGAGTACCTCAT  
CCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA  
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA  
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTGAGAAGGAAGGATG  
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT  
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC  
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT  
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTTGCTTGTGGAAAGACTG  
TTTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT  
GATTACCTCTGGTGTGACAGGTTTGAAGTTGCACTTCTTAAGGAACAGCCATAATCCTCTG  
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTGTTATAGGAAGTTGTA  
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC  
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC  
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT  
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA  
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT  
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTTATTTGTAAGACATTACTTATTAAG  
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG  
TACTACAGATTTTCAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTTCTTTAGT  
GCAATACAATAAACTCTGAAATTAAGACTC

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
```

&lt;subunit 1 of 1, 144 aa, 1 stop

&lt;MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAF  
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA  
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

## **FIGURE 130**

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG  
AATCCCCTTGTA TACTCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC  
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCCTTTGGCATATCATATTTGGAGGTATA  
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT  
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT  
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT  
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT  
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

1002033366

## FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG  
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC  
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGGACCGCTTCGGCTGAAGCATTTGAC  
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACAC  
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC  
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA  
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC  
ATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC  
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC  
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC  
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA  
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA  
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACCTCT  
TGTCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC  
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGCTGACTTGGAGTTTATGAATGAACAAAAG  
CTAAACAGATATCCAGCTTCTTCTTGTGGTTGTTAGATCTAAAACCTGAAGATCATGAAGA  
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT  
AAAAGACAAGTGTAATAGACATCTAAAATTCCACTCCTCATAGAGCTTTTAAAATGGTTTCA  
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG

## **FIGURE 132**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQCLTYPLHTYP  
KEEELYACQRCRLFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLGCQNQLPFA  
ELRQEQLMSLMPKMHLFPLTLVRSFWSMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI  
QYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGLRCLSLNSGWILTTTLVL  
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG  
PLPTKVNLAHSEI

### **Important features:**

#### **Signal peptide:**

amino acids 1-31

#### **Transmembrane domain:**

amino acids 241-260

#### **N-glycosylation site.**

amino acids 90-93

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACACCTACCC  
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG  
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA  
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCGC  
TGAAGTGAAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA  
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

400520 535466

FIGURE 134

[illegible]

## FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCCGGAGGT  
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGCGACCGAGC  
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGGCCGCGGCTG  
GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC  
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT  
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA  
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT  
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA  
GTTCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA  
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA  
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC  
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA  
CAAGGGACCAGATGCTTGGAATAATGGAATGTCTACGAAGAAAACCTGTTTAAAGCCAC  
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT  
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG  
CCTACATGCAAGCATTAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG  
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA  
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAATAAGGGCTTT  
ATCCAAAGTGTTACCATTCTTCGAGCGCCCAGATTTTCAACTCTTTACTGGAAATAAAATTC  
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTG  
CATTTTGATGAGAATTCATTTTTTGGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA  
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC  
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG  
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCATCTAACCAGACAAGA  
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAAC  
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC  
AATGGAGGCGAAAGAGTGAATTTTATTCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA  
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT  
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG  
GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 136**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAQAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF  
PRLQKLLSDYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA  
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNP  
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAF  
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI  
ELRALSKVLPFFERPDLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH  
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH  
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **N-glycosylation site.**

amino acids 280-283 and 384-387

#### **Amidation site.**

amino acids 94-97

#### **Glycosaminoglycan attachment site.**

amino acids 20-23 and 223-226

#### **Aminotransferases class-V pyridoxal-phosphate**

amino acids 216-222

#### **Interleukin-7 proteins**

amino acids 338-343

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

*[The page contains extremely faint, illegible text.]*

## FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA  
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC  
CGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTGCCCCCTGGTGTTCCTGACAGG  
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG  
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC  
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG  
GGCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC  
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC  
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA  
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCTAGAAAGCAAA  
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT  
GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCCCAGGCCTGTGCCCCCTCTCTGGTCT  
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA  
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT  
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG  
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCCTAACAGA  
TTCAGACTCCTGGCCAGGTGTGGTGGCCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG  
GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT  
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA  
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG  
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA  
AATAATAATAATAATAAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA  
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG  
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

## **FIGURE 139**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLPFGPPEAEFGYSVLQHVGGRWMLVGAPW  
DGPSGDRRGDVYRCPVGGAHNAPCAKGLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **Cell attachment sequence.**

amino acids 70-73

#### **N-glycosylation site.**

amino acids 98-101

#### **Integrins alpha chain proteins**

amino acids 67-81

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

[illegible]

## FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG  
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA  
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT  
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA  
**AATG**CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT  
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC  
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA  
AACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT  
GGATCCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC  
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG  
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA  
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTTGAGTTC  
CTTGTTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG  
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA  
CATTCTGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA  
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT  
GGTCTGTGCCACTGTTCTGTCTGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG  
TGGTCCTCCCAGACACCTTGAAAATAACCAATTACCCCAAGTTAATCAGCTGCAGAAGG  
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT  
CTCA**TAG**GTTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC  
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA  
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC  
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC  
CTGGGAAAAGTGACTTCATCCCTTCGGTCTTAAGTTTTCTCATCTGTAATGGGGGAATTACC  
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA  
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTGAG  
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT  
GGCTTGAGAGAGCCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG  
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT  
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC  
CATGAACACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT  
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTTC  
TGTTGGTAAAGTACAGAATTGAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA  
AAAAAAAAA

094655 07004

## **FIGURE 142**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGE  
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS  
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG  
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMILILV  
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

### **Important features:**

#### **Signal peptide:**

amino acids 1-29

#### **Transmembrane domain:**

amino acids 230-255

#### **N-glycosylation site.**

amino acids 40-43 and 134-137

#### **Tissue factor proteins.**

amino acids 92-119

#### **Integrins alpha chain proteins**

amino acids 232-262

## **FIGURE 143**

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT  
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA  
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG  
TCAAAC TGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT  
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC  
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA  
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCT  
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG  
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAAGGGCCACATTGGGC  
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC  
CCGACCTGGGATGGAGATCACCAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG  
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG  
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC  
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT

## FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA  
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA  
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG  
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC  
GGAGAGGAGGTGTGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA  
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG  
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG  
GCTGCTTTGGCATTGTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT  
GGAGGGACAGGGTCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGACAGGGCGTTGGG  
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC  
GCTGGTACTCTGGGCTGCACTGGGGGACAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG  
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC  
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA  
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCATTAAGGCTCAGCACTGGAGGAGAGA  
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT  
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAAGTGCAGGCTGCT  
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG  
AGGTGCAGCTCATTCACTTCAACCAGGAACTCTACGGGAATTTACGCGCTGCCTCCCGCGGC  
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTC  
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC  
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC  
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA  
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT  
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC  
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGA  
TGGTGTCCCCCATGGTCGCTGAGACTCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTC  
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

## **FIGURE 145**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC  
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL  
LYSHRLSELRLFLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL  
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE  
TVTWILIDRALNITSLQMHSRLRLSQNPFSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER  
RCRGPNYRLHVDGVPHGR

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 177-199

#### **N-glycosylation site.**

amino acids 118-121, 170-173 and 260-263

#### **Eukaryotic-type carbonic anhydrases proteins**

amino acids 222-270, 128-164 and 45-92

## FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC  
GTTCTGTGTCCCGCCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAA  
GCAGCTCCCTTCCCACCCCACTGCAGGTCTAATTTTGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGC  
CGCAGAGGCGGAGGCTCGCGTATTCTGCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTC  
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCTGGTTTGGCT  
CACCTCTCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA  
ATCCTGAGGTCAATTCATTATGAAGTGTACCGCGCGGGAGTGGCTCAGAGTAACACAGTGCTGTTTCATGGCTAGA  
GCAATTCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT  
GAGTGGTGGATAGCCAAACAACGAGGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT  
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAGA  
TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT  
TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT  
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTAGGTGTTCTGGCCCTGTATGTACACATTATACA  
CAGGTCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTACATAACATGAACATCTGGGGGAG  
ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAACTGGTGGGGCCATCCCCCTTACAAA  
CATGGGCGGCCCTGTTCTGCTTGGCCACCTAGTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG  
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC  
CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCT  
TGTGAAGTAAGATTAAAGAGTCAGTGCAAGGAACAACCTGCAATAGGTACGAATGTCTGCTGGCTGTTTGGAT  
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT  
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAATAGA  
AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTACAGTCTCTAAAGTAACAGTTCAGGCTGTG  
ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCT  
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGAATTGGAACCTCGAGTTTATTCTGATCTGTCCAGTATC  
TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCCTGTGGACAAAAGA  
AAGACCTACATTGCTTCTTTTTCAGAAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT  
AGAGTGTGTTGCTGTTGTGTGAAGTGAATACTTGAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA  
ATTTTGTATAAACTGTAACTTACTGTACAGAGTACATCAACTATTTTCAGCCCAAAAAGGTGCCAAATGCATA  
TAAATCTTGATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG  
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT  
GTTCTACGTTTCATATATTATATGGTGCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG  
GCCCTCAGAAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACTGAAAGAAACCTTATCACATTTTCCCC  
AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC  
TGTTAATTTTAGGCATATAGAATATTAAATTTCTGATATGCACTTCTTATTTTATATAAAATAATCCTTTAATATC  
CAAATGAATCTGTTAAATGTTTGATTCCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT  
ATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA  
AAATTGAGGTACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT  
GAACAAAGATGAACTAATGTATTACATTACCATTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA  
ATATTGCCATATCATGGTACCTATAATGGTGATATATTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA  
AATCTGTAAATGTTAGTTTTTGGTAATTTTTTTTCTGCTGCTGGATTACATATTAAATTTTTTTCTGCTGGTGA  
TAAACATTAAATTAATCATGTTTCAAAAAAAAAAAAAA

## **FIGURE 147**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLLEKYMDEDEGEWWIAKQRGKRAITDNDM  
QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW  
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLC  
HNMNIWGQIWPKAVYLV CNYSPKGNWWGHAPYKHGRPC SACPPSFGGGCRENLCYKEGSDRY  
YPPREEETNEIERQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR  
YEC PAGCLDSKAKVIGSVHYEQSSICRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI  
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG  
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRV  
FAVV

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein**

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

#### **N-glycosylation site**

amino acids 28-31

## FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG  
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT  
GGCGTCTCCGGGCGCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCG  
CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT  
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC  
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT  
CATCACCAAGAGCGTGTCCAGGTACCGAGGTGAGGAGCACTGCCTGCACCCCCAAGCTGC  
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA  
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA  
CTTTGCAGATTAaaAGCCTTTC  
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTTACCAACGGTCAG  
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT  
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTTAA  
AAAATGCTTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGGA  
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTTGCTCCATTCCCTAGCTTGGGAAGC  
TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG  
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT  
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT  
GGTTAATTTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC  
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA  
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT  
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA  
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT  
TTGTGATTCCCCATGTAATTCTTCAATGTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT  
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC  
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT  
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA  
AAAAAAAAAA

## FIGURE 149

MSLLPRRAPPVSMRLAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH  
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

**Signal sequence:**

amino acids 1-34

MSLLPRRAPPVSMRLAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH  
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

[illegible]

GCCCAAGGACTGCTATGGCTTCTTTTGTGTTTACCCCGGTCTGCGT**CATGTTAA**ACTCCAATGTCTCTGTG  
 GTTAACTGCTCTTGCCATCAAGTTCACCCCTCATTTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG  
 CAAAATCCGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCACTACTTAGGGGTCCCCTA  
 TGCTTACCCCCCACTGGAGAGAGGCGGTTTTCAGCCCCCAGAACCCCCGTCTCTCTGACTGGCATCCGAAATAC  
 TACTCAGTTTGTGCTGTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT  
 TACCGCCAATTTTGGACTACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTAAACACTTCAGT  
 GCCACGGGAAGATGGAGCCCAACAAAGAAAAACGCAGATGATATAACAGAGTAATGACCGTGGTGAAGACGAAGA  
 TATTTCATGATCAGAACAGTAAGAAGCCCGTTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCA  
 CATGATTGACGGCAGCATTTTGGCAAGCTACGGAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATAC  
 AGGGTTTTTTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG  
 GATTGAGGAGAAATGTGGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCCT  
 CTGTGTACGCCTGTGACCTGTCCCACCTACTCAGAAGTCTCTTCCAGAAGGCCATCATTCAGAGCGGCACCGC  
 CCTGTCCAGCTGGGCAGTGAACCTACCAGCCGGCCAAAGTACACTCGGATATTGGCAGACAAGGTCGGCTGCAACAT  
 GCTGGACACCACGGACATGGTAGAATGCCGTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC  
 GGCCACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGACGTCATCCAGACGACCCCCAGATCCTGATGGA  
 GCAAGGCGAGTTCCCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCCTGAAGTTTCGTGGACGGCAT  
 CGTGGATAACGAGGACGGTGTGACGGCCCCAACGACTTTGACTTCTCCGTGTCCAACCTCTCGTGGACAACCTTTACGG  
 CTACCTCTGAAGGGAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAAACCC  
 GGAGACGCGCGGAAAAACCTTGGTGGCTCTCTTTACTGACCACAGTGGGTGGCCCCGCGCGTGGCCGCCGACCT  
 GCACGCGCAGTACGGCTCCCCACCTACTTCTATGCCCTTCTATCATCACTGCCAAAGCGAAATGAAGCCAGCTG  
 GGCAGATTTCGGCCCATGGTGATGAGGTCCCTATGTCTTCGGCATCCCATGATCGGTCCCACCGAGCTCTTCAG  
 TTGTAACTTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTTCATGACCCTACTGGACGAACTTCGCCAAAACCTG  
 TGATCCAAATCAACCAGTTCTCTCAGGATACCAAGTTTCATTACACAAAACCCACCGCTTTGAAGAAGTGGCCTG  
 GTCCAAGATATAAATCCCAAGACGAGCTCTATCTGCATATTGGCTTGAACCCAGAGTGAGAGATCACTACCGGGC  
 AACCAAGTGTGCTTTCTGGTTTGGAACTCGTTCTCATTTTGACAACTTTGAACGAGATATTCAGTATGTTTCAAC  
 AACCACAAAGGTTCTTCCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC  
 AACCACCAAACGCCCAGCAATCACTCCTGCCAACAAATCCCAAACACTCTAAGGACCCCTCACAAAACAGGGCCTGA  
 GGACACAACCTGTCTTCATTGAAACCAAACAGAGATTATTCACCAGAAATTAAGTGTCAACATTGGCCGTGGGGCGCTG  
 GCTCCTCTTCTCCTCAACATCTTAGCTTTTGGCGCGCTGACTACAAAAAGGACAAGAGGCGCCATGAGACTCAG  
 GCGCCCCAGTCCCCAGAGAAACACCAAAATGATATCGCTCATCTCCAGAACAGAGAGATCATGTCTCTGCAGAT  
 GAAGCAGCTGGAACACGATCACGAGTGTGAGTTCGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGA  
 CTACACCCCTCACGCTGCGCCGGTGCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAA  
 CACACTGACGGGGATGCAGCCTTTGCACACTTTTAAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA  
 CGGACATTTCCACCACCTAGAGTA**TAG**CCTTTGCCCTATTTCCCTTCTCATCCCTCTGCCCTACCCGCTCAGCAACAT  
 AGAAGAGGGAAGGAAGAGAGAAGGAAGAGAGAGAAGAAAGTCTCCAGACAGGAATGTTTTGTCTCCACT  
 GACTTTAAGACAAAAATGCAAAAAGGCAGTCACTCCCATCCCGGCAGACCCTTATCGTTGGTGTTTTTCCAGTATTAC  
 AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGTTAAAATAACTGCTTTAAGATCTCTACCA  
 CTCCAATCAATGTTTTAGTGTGATAGGACATCACCATTTCAAGGCCCGGGTGTTCCTAACGTCATGGAAGCAGCT  
 GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAAATTAATGGAAGTTTAAACATTTCTTTCTGTGC  
 CACACAATGGATGGCTCTCCTTAAAGTGAAGAAGAGTCAATGAGATTTTGCCAGCACATGGAGCTGTAATCCAG  
 AGAGAAGGAACGTAGAAATTTATTATTATAAAGAATGGACTGTGCAGCGAAATCTGTACGGTCTGTGCAAGAG  
 GTGTTTTTGCCAGCCTGAACATATATTTAAGAGACTTTGT

姓名	性别	年龄	籍贯	职业	文化程度	政治面貌	健康状况	婚姻状况	子女情况	其他
王德胜	男	45	山东	工人	高中	党员	良好	已婚	2子1女	
李秀英	女	42	河北	教师	大学	党员	良好	已婚	1子1女	
张国强	男	38	河南	农民	初中	团员	良好	已婚	2子1女	
刘小红	女	35	江苏	医生	大学	党员	良好	已婚	1子1女	
陈为民	男	50	浙江	干部	高中	党员	良好	已婚	2子1女	
赵子龙	男	48	湖北	工人	初中	团员	良好	已婚	2子1女	
周美兰	女	40	湖南	教师	大学	党员	良好	已婚	1子1女	
吴大伟	男	36	四川	农民	高中	团员	良好	已婚	2子1女	
孙丽娟	女	33	广东	医生	大学	党员	良好	已婚	1子1女	
郑为民	男	47	广西	干部	高中	党员	良好	已婚	2子1女	
冯小华	女	39	福建	教师	大学	党员	良好	已婚	1子1女	
马国强	男	44	江西	工人	初中	团员	良好	已婚	2子1女	
朱小红	女	37	山西	医生	大学	党员	良好	已婚	1子1女	
徐为民	男	49	安徽	干部	高中	党员	良好	已婚	2子1女	
郭小华	女	41	陕西	教师	大学	党员	良好	已婚	1子1女	
何国强	男	34	甘肃	农民	初中	团员	良好	已婚	2子1女	
李丽娟	女	32	宁夏	医生	大学	党员	良好	已婚	1子1女	
周为民	男	46	青海	干部	高中	党员	良好	已婚	2子1女	
吴小华	女	38	新疆	教师	大学	党员	良好	已婚	1子1女	
马国强	男	43	内蒙古	工人	初中	团员	良好	已婚	2子1女	
朱小红	女	36	吉林	医生	大学	党员	良好	已婚	1子1女	
徐为民	男	48	辽宁	干部	高中	党员	良好	已婚	2子1女	
郭小华	女	40	黑龙江	教师	大学	党员	良好	已婚	1子1女	
何国强	男	35	河北	农民	初中	团员	良好	已婚	2子1女	
李丽娟	女	33	山东	医生	大学	党员	良好	已婚	1子1女	
周为民	男	47	河南	干部	高中	党员	良好	已婚	2子1女	
吴小华	女	39	江苏	教师	大学	党员	良好	已婚	1子1女	
马国强	男	44	浙江	工人	初中	团员	良好	已婚	2子1女	
朱小红	女	37	湖北	医生	大学	党员	良好	已婚	1子1女	
徐为民	男	49	湖南	干部	高中	党员	良好	已婚	2子1女	
郭小华	女	41	四川	教师	大学	党员	良好	已婚	1子1女	
何国强	男	34	广东	农民	初中	团员	良好	已婚	2子1女	
李丽娟	女	32	广西	医生	大学	党员	良好	已婚	1子1女	
周为民	男	46	福建	干部	高中	党员	良好	已婚	2子1女	
吴小华	女	38	江西	教师	大学	党员	良好	已婚	1子1女	
马国强	男	43	山西	工人	初中	团员	良好	已婚	2子1女	
朱小红	女	36	安徽	医生	大学	党员	良好	已婚	1子1女	
徐为民	男	48	陕西	干部	高中	党员	良好	已婚	2子1女	
郭小华	女	40	甘肃	教师	大学	党员	良好	已婚	1子1女	
何国强	男	35	宁夏	农民	初中	团员	良好	已婚	2子1女	
李丽娟	女	33	青海	医生	大学	党员	良好	已婚	1子1女	
周为民	男	47	新疆	干部	高中	党员	良好	已婚	2子1女	
吴小华	女	39	内蒙古	教师	大学	党员	良好	已婚	1子1女	
马国强	男	44	吉林	工人	初中	团员	良好	已婚	2子1女	
朱小红	女	37	辽宁	医生	大学	党员	良好	已婚	1子1女	
徐为民	男	49	黑龙江	干部	高中	党员	良好</			

Signal sequence:

Transmembrane domains:

amino acids 189-204, 675-692

## FIGURE 152

GGGAAAGATGCGGCGGACTCTGGGACCCCTTGGGTCTGTGGCAGCAGTGGCGGCGATGTTTGT  
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCA  
CAGCAAGTCGGGGCGGGTCAAACGTTTCGAGTACTTGAAACGGGAGCACTCGCTGTGGAAGCC  
CTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA  
TGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG  
GTGCCATGTTTCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAA  
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC  
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT  
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT  
CAGCTATGATCATGAGCGGGATGGGCGGCCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC  
GCAATCTTCATTACGACACCTTCTCTGGTGATTGCTACGTCAAGAGGCATTTGACGATAATG  
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC  
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA  
TTTCCTTGAAGTTGTTTGAAGTACAGTGGAGAGAACCCAGAAAGAGGAAAAGCTCCATCGA  
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC  
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGTGTCTTCTGTATTTGCCA  
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA  
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG  
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG  
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT  
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCTGCAGTCC  
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTGTGTGATGCCAAAATCACAGAAC  
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGAAGTCAAGAGGCCCTTCTACTTCAGTTTTG  
AATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCTGTTGGGTT  
TTGCATTTGACCCAACCTCTGCCTACCTGAGGAGCTTTCTTTGAAACCAGGATGGAACT  
TCTTCCCTGCCTTACCTTCTTTCACTCCATTCAATTGTCTCTCTGTGTGCAACCTGAGCTG  
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTAC  
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT  
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG  
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG  
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTATGTG  
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT  
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT  
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT  
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT  
CACTGTCAATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA  
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG  
AGTCTCCTTGGAGGGCCTGGAACCTCTGAGTCTCCTATGAACCTCTGTAGCCTAAATGAAAT  
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG  
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

[illegible]

```
><subunit 1 of 1, 348 aa, 1 stop
```

MAATLGPLGSWQQWRRCLSDRGSRMLLLLLLLLLGSGQGPQQVGAGQTFEYLRKREHSLSKPYQ  
GVGTGSSSLWNLMGNAVMVTQYIRLTQDMQSKQALWNRVPCFLRDWELQVHFKIHGQGKKN  
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTPNEEKQQERVFPIISAMVNNGSLSY  
DHERDGRPTLGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG  
YYFGTSSITGDLSDNHVISLKLFEFTVERTPEEEKLHRDVFLPSVDNMKLPENTAPLPPLS  
GLALFLIVFFSLVFSVFVFAIVIGIILYNKWQEQSRKRFY

amino acids 1-38

## amino acids 310-329

**SECRET**

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATTGGGCGTGAGTGCAATCTACGGATCAGTCTCTGATGGTGGGTCTGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCTGACATGTTCCCGATTGAGGTAAGAACATGAAGAGAAAATAGAATACTTAATAATGCTCTTTCCGCAACCGCTTCTTGCTGCTGTGGCCCTGGCTGCCTGCTGGCTTTGTGAGCCTCAGCCTGCAGTCTTCCACCTGATCCCGGTGTGCATCTCTAAGAATGGAATGAGTAGCAAGAGCTCGAAGAGAAATCATGCCGACCCCTGTGACGGAGCCCCCTGTGACAGCCCCGTTTTATGAAGCTCTTTTGTACTGCAACATCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCATGCCCGCATCATTTTTAAGCTGGTCTCAGTGCATGTGTTTCATTGCCACGGAGACAGGTACCCACTGTATGTCATTCCCCAAACAAAGCGACGAGAAATGACTGCACCTGTTGGCTAACAGGAAACCGTATACCCAAAACTGGAAGCTTTCA TTAGTCACATGTCAAAAGGATCCGGAGCCTCTTTGCAAGAGCCCTTGAACCTCTTGCCCTCTTTACCCAAATCACC CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTGTGAGAAGCGGTACGTTGCTGTGGGGATA TCTATCTAAAGAAAACAAACTCCTGCCCAATGATTGGTCTGCAGACATTAGCTCTATTTAGAGACCACTGGGAAAA GCCGGACCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCAGATTTTGACTGGAAGAAGATTTATTTCA GGCACAGCCAAAGTGCCTGTTCTGCTCTGGAAGCTGCTATTGCCCGGTGAAGAAACAGTATCTGGAAAAGGAGC AGCGTCGTACAGTACCTCTACGTTTGA AAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAAGATCGTGG ATGTCCCCCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCAGTCTGCTGCCACTTCTGCCCAAGATGACGT TTCCCTGTACCAAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA GGGAAAGACGGGAGAAGAAATTTGACTTCCGGGTATTCTCTCCTGGGTGCCCAACCCATCCTGAACCAAAACCATCG GCCGATGTCAGCGTGCACCGGAGGGCAGGAAAGAAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACCTCTGT CACCAGTCTCAGTGCCTTTGGGCCTTTCAGAAGCCAGGTTCCCAAGGTTTGAGCCAGGTTGATCTTTGAGCTTT GGCAAGACAGCAAGAAAGCCCACTGAACATTCGTCCGGATTCTTTCAAGTGGCGTCGATGTCACATTTCCACACCC CT'TTCTGCCAAGACCACCACAAGGCTTCTCCCAAGCCCATGTGCCCGCTTGAAAACTTGGTCCGCTTTGTGAAAA GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAAT'TATTATGATGCATGTCACAGGGAAGGATTCTAAAGG TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC TAAGGGTAGAAGATTATTGCTTTTTTAAAGGCTAAATATTGTTTGTGGGAACCAAGATGGTTGGGGTTGAACAGT AAGCATATTGCTGCAATGTGGTACGTGAATTTGCTTGGTACAAAAATGGCCAGTTACAGAGGAATAGAAGGTACTT TATCATAGCCAGACTTCCGTTAGAATTGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGAC TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAAATTTTTTTTAATCTTAGACATT TTTACCTTGTCCTTGTTAAGAATTTCTTGAAGTGATTATCTAAATAAAGGTTGGCAAACTTTTTCTGTAAAGG GCCAGATTGTAAATATTTACAGACTGTGTGGACCAAAGGCCACATACAGTCTCTGTCTATAACTCAACTCATCTGT TTTCTGAAGCAGGAAAGCCACACAGCATACATAAGGAATATGTGTAGCTGGTTCGAGGCCAGACAAAACA GATGTGACCAGCACTTGGCCCCGTGGGCTGTAGTTTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCCTCCAGTAACCTTCTGCTAGAAACA CAGAAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTTTAAGTATGTTCTAAATATTTGT CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTATAGTTTAGGAAAATATT TTCTAAGACAGGTTTTAGATGATCTTATTCTGTAGTAAATTTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG GAGGCTAGAAGATGAATTCAGGCACCTTTCTTCCAATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA ACTGGAATCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATCTCGATTGATTTTTTAAATGCGTTTTTGA AGAACCTTGCTATTAGGTAGTTTTACAGATCTTTATAAGGTGTTTTATATATAGAAGCAATTTAATTTACATCTG TGATTTCTGAACATAAATGGTGCTAATTACAGAAATGGAAGTGAAGTGAAGTATCTCTGTTGTCATCGGCATCC AACTTTTTCTCTTTGTTTTTGTCCAGTTGTGCATTGGAATATGTCTGTTTTCTATAAATAAAATTTTTTAAGAATA

姓名	性别	年龄	职业	住址	电话	备注
王德胜	男	45	工人	XX路XX号	XXXX	
李小红	女	32	教师	XX街XX号	XXXX	
张小明	男	28	学生	XX巷XX号	XXXX	
赵大伟	男	55	干部	XX里XX号	XXXX	
陈丽娟	女	40	护士	XX院XX号	XXXX	
周国强	男	38	工程师	XX厂XX号	XXXX	
吴小芳	女	25	职员	XX楼XX号	XXXX	
郑大刚	男	60	农民	XX村XX号	XXXX	
孙文娟	女	35	医生	XX诊所	XXXX	
马小强	男	22	学生	XX校XX号	XXXX	
徐大伟	男	50	工人	XX厂XX号	XXXX	
黄小娟	女	30	教师	XX校XX号	XXXX	
周国强	男	48	干部	XX局XX号	XXXX	
吴小芳	女	28	学生	XX校XX号	XXXX	
郑大刚	男	58	工人	XX厂XX号	XXXX	
孙文娟	女	33	护士	XX院XX号	XXXX	
马小强	男	23	学生	XX校XX号	XXXX	
徐大伟	男	51	工人	XX厂XX号	XXXX	
黄小娟	女	31	教师	XX校XX号	XXXX	
周国强	男	49	干部	XX局XX号	XXXX	
吴小芳	女	29	学生	XX校XX号	XXXX	
郑大刚	男	59	工人	XX厂XX号	XXXX	
孙文娟	女	34	护士	XX院XX号	XXXX	
马小强	男	24	学生	XX校XX号	XXXX	
徐大伟	男	52	工人	XX厂XX号	XXXX	
黄小娟	女	32	教师	XX校XX号	XXXX	
周国强	男	50	干部	XX局XX号	XXXX	
吴小芳	女	30	学生	XX校XX号	XXXX	
郑大刚	男	60	工人	XX厂XX号	XXXX	
孙文娟	女	35	护士	XX院XX号	XXXX	
马小强	男	25	学生	XX校XX号	XXXX	
徐大伟	男	53	工人	XX厂XX号	XXXX	
黄小娟	女	33	教师	XX校XX号	XXXX	
周国强	男	51	干部	XX局XX号	XXXX	
吴小芳	女	31	学生	XX校XX号	XXXX	
郑大刚	男	61	工人	XX厂XX号	XXXX	
孙文娟	女	36	护士	XX院XX号	XXXX	
马小强	男	26	学生	XX校XX号	XXXX	
徐大伟	男	54	工人	XX厂XX号	XXXX	
黄小娟	女	34	教师	XX校XX号	XXXX	
周国强	男	52	干部	XX局XX号	XXXX	
吴小芳	女	32	学生	XX校XX号	XXXX	
郑大刚	男	62	工人	XX厂XX号	XXXX	
孙文娟	女	37	护士	XX院XX号	XXXX	
马小强	男	27	学生	XX校XX号	XXXX	
徐大伟	男	55	工人	XX厂XX号	XXXX	
黄小娟	女	35	教师	XX校XX号	XXXX	
周国强	男	53	干部	XX局XX号	XXXX	
吴小芳	女	33	学生	XX校XX号	XXXX	
郑大刚	男	63	工人	XX厂XX号	XXXX	
孙文娟	女	38	护士	XX院XX号	XXXX	
马小强	男	28	学生	XX校XX号	XXXX	
徐大伟	男	56	工人	XX厂XX号	XXXX	
黄小娟	女	36	教师	XX校XX号	XXXX	
周国强	男	54	干部	XX局XX号	XXXX	
吴小芳	女	34	学生	XX校XX号	XXXX	
郑大刚	男	64	工人	XX厂XX号	XXXX	
孙文娟	女	39	护士	XX院XX号	XXXX	
马小强	男	29	学生	XX校XX号	XXXX	
徐大伟	男	57	工人	XX厂XX号	XXXX	
黄小娟	女	37	教师	XX校XX号	XXXX	
周国强	男	55	干部	XX局XX号	XXXX	
吴小芳	女	35	学生	XX校XX号	XXXX	
郑大刚	男	65	工人	XX厂XX号	XXXX	
孙文娟	女	40	护士	XX院XX号	XXXX	
马小强	男	30	学生	XX校XX号	XXXX	
徐大伟	男	58	工人	XX厂XX号	XXXX	
黄小娟	女					

><MW: 55240, pI: 9.30, NX(S/T): 2

amino acids 1-18

## FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGAACAGTACTGTGGA  
GCTATTTAAGAGATAAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC  
GCCGCTGTTACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGACGCGCCTGAAGCACAAAGCAGAT  
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTAT  
GGGCGAAGGAACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAATGAT  
TCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG  
ATGCACCCAGATACGCTATTCTGATTCGGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT  
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCCCCAGGGTAGGACGCAGCTTTTCGCCCT  
GAATCCGCGCAGCGGCAGCTTGCTCACGGCGGGCAGGATAGACATCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG  
TCAATTAATCTAGACATCTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA  
CGACAATGCGCCTTACTTTCTGTGAAAGTGAAATTAGAAATAAAAATTAGTGAAAATGCAGCCACTGAGATGCGGTT  
CCCTCTACCCACGCTTGGGATCCGATATCGGGAAGAAGCTCTCTGCAGAGCTACGAGCTCAGCCCCGAACACTCA  
CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCCCGAATTGGTGCTGAAACGCGCCCTGGACCG  
CGAAGAAAAGGCTGCTCACCACCTGGTCCTTACGGCCTCCGACGGGGGCGACCCGGTGCGCACAGGCACCGCGTCC  
CATCCGCTGATGGTTCTGGATGCGAACGACAACGCACCGAGCTTTGCTCAGCCCGAGTACCGCGCGAGCTTCC  
GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGCTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGCGGAAGT  
GAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCCAAGTTTTCAAAGTAGATTGTAATTGAGGACAATATC  
AACAAATAGGGGAGTTGGACCACGAGGAGTCAGGATTTCTACCAGATGGAAGTGCAAGCAATGGATAATGCAGGATA  
TTCTGCGCGAGCCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCTCACCTCTCT  
CGCCAGCTCGGTTCCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA  
GGAAAACGGACAGGTGATCTGTTTCAATCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTACTA  
TAGTTTATGTCACAGACATAGTCTTGGATAGGGAACAGGTTCCCTAGCTACAACATCACAGTGACCGCCACTGACCG  
GGGAACCCCGCCCCCTATCCACGGAACTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGCGGTCTT  
CCCTCAGGCCTCCTATTCGCTTATATCCAGAGAACAAATCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA  
CGACCCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC  
GTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA  
CTTGCAAGTGAAAGTGATGGCGCGGGACAACGGGCACCCGCCCTCAGCAGCAACGTGTCTGTTGAGCCTGTTCTG  
GCTGGACCCAGAACGACAATGCGCCCCGAGATCCTGTACCCCCGCCCTCCCCACGGACGGTTCCACTGGCGTGAGCT  
GGCTCCCCGCTCCGACAGAGCCCGCTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC  
CTGGCTGTCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG  
CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGCCGTCCAGGACCACGGCCAGCC  
CCCTCTCTCCGCCACTGTACGCTCACCCTGGCCGTGGCCGACAGCATCCCCAAGTCCCTGGCGGACCTCGGCAG  
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCTG  
CGTCTTCTGGCCTTCTGTCATCTTGCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCCTGCTGCAGGC  
TTCAGGAGGCGGCTTGACAGGAGCGCCGGCTCGCACTTTGTGGGCGTGACGGGCTGCAGGCTTTCTGCGAGAC  
CTATTTCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACTGATCTTCCCCAGCCCCAACTATGCAGA  
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAAGCGAGCCCCCTTTTGCTGTGAGGTGATTCCGTATTTTCTAAAGA  
CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTTCTTTTAAATTGCTCTGTCTCCCAAGC  
TGGAGTGACGCGTACGATCATAGCTCACTGCGGCCTCAAACCTCTAGGCTCAAGCAATTATCCACCTTTGCCT  
CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTATCTAT  
CTATCTATCTATCTATCTATTACTTTCTTGACAGACGGGAGTCTCACGCCTGTAATCCAGTACTTTGGGAGGC  
CGAGGCGGGTGGATCACCTGAGGTGGGAGTTTGAGACCAGCTGAGCAACATGGAGAAACCCCGTCTATACTAA  
AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT  
TGCTTTAACCTGGAGGTGGAGGTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG  
AAACTCTATCTCA

## **FIGURE 157**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER  
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR  
DINDNAPYFRESELEIKISENAATEMRFPPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA  
DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR  
ASVPENLALGTQLLVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGTISTIGELDHE  
ESGFYQMEVQAMDNAGYSARAKVLIITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLVN  
DQDSEENGQVICFIQGNLPFKLEKSYGNYSLVTDIVLDREQVPSYNITVTATDRGTPPLST  
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI  
QGASLSSYVSINSDTGVLIALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN  
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG  
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLESPA  
NSETSDLTLYLVVAVAAVSCVFLAFVILLALLALRLRRWHKSRLQASGGGLTGAPASHFVGVD  
GVQAFLQYTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL  
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYZRCKPPTVCLS  
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

### **Signal sequence:**

amino acids 1-30

### **Transmembrane domains:**

amino acids 693-711, 809-823, 869-888

## FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG  
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA  
TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT  
GCGGCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC  
ACAGGAGGACAAGGTGCTGGGGGGTCA TGAGTGCCAACCCATTTCGCAGCCTTGGCAGGCGG  
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTG TAGGTGGCAACTGGGTCTT  
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA  
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTT CAGTCCATCCCACACCCCTGCTACAACA  
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC  
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG  
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT  
GTGCAGAAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA  
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG  
CCCCCTGGTGTGTGATGGTGC ACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA  
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC  
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACA ACTCT  
CTGGTTC

## **FIGURE 159**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL  
VGGNWWLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIHPHCYNSSDVEDHNHDLMLL  
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENF PDTLNCAEVKIFPQKKCED  
AYPGQITDGMVCAGSSKGADTCQGDGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY  
LDWIKKIIGSKG

### **Important Features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 51-71

#### **N-glycosylation site.**

amino acids 110-113

#### **Serine proteases, trypsin family, histidine active site.**

amino acids 69-74 and 207-217

#### **Tyrosine kinase phosphorylation site.**

amino acids 182-188

#### **Kringle domain proteins motif**

amino acids 205-217

## FIGURE 160

GGCGCCGGTGACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCCG  
CGCGCCGCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCCGCGCCCCG  
GCCCCGCGCCAGGTGAGCGCTCCGCCCCGCGCGAGGCCCGCCCCGGCCCCCGCCCCG  
CCCCGGCCGGCGGGGGAACCGGGCGGATTCTCTGCGCGTCAAACCACCTGATCCCATAAAAC  
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCCGCCCCCTCG  
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGC  
GGAGCGCGCCGAGCCTCGTCCCGCGGCCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA  
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCCGCCCGAAACGACTTTTCAGTCCCCGACGCGC  
CCCCCCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG  
CTGTGGCTGCGAGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCTGCGTATGCTACAATGA  
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCCGTGGGCATCCCTG  
CTGCCAGCCAGCGCATCTTCTGACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC  
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGC  
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC  
GGTCTGTGGACCTGCCACATTCCACGGCCTGGGCCGCTACACACGCTGCACCTGGACCGC  
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA  
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA  
CACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCCGAGCGCGCCTTCCGTGGGCTG  
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTT  
CCGTGACCTTGGCCGCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA  
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG  
TGTGACTGCCGGGCACGCCCCACTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA  
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG  
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC  
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT  
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCCGGTG  
ACAGCCCGCCGGGCAACGGCTCTGGCCCACGGCACATCAATGACTCACCCCTTTGGGACTCTG  
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT  
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCCGCAGCCACT  
GCCGTCTGGGCCAGGCAGGCAGCGGGGTGGCGGACTGGTGAAGCTCAGAAGGCTCAGGTGCC  
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT  
TGGGCCCTGCTGAACCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC  
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG  
GTCCTCCCTGATGGACGCTGCCGCCCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG  
TTCGGCGGCAGCGTTTGTTCAGAACGCCGCTCCACCCAGATCGCGGTATATAGAGATAT  
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA  
AAAA

## **FIGURE 161**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI  
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA  
TFHGLGRLHTLHLDRCLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH  
GNRISSVPERAFRGLHSLDRLLHLHQN RVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP  
LRALQYLR LNDNPWVCD CRARPLWAWLQKFRGSSSEVP CSLPQRLAGRDLKRLAANDLQGCA  
VATGPYHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGN  
GSGPRHINDSPFGTLP GSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA  
GSGGGGTGDSEGS GALPSLTCSLTPLGLALVLWTVLGPC

### **Important features:**

#### **Signal peptide:**

amino acids 1-26

#### **Leucine zipper pattern.**

amino acids 135-156

#### **Glycosaminoglycan attachment site.**

amino acids 436-439

#### **N-glycosylation site.**

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

#### **VWFC domain**

amino acids 411-425

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

[illegible]

## **FIGURE 163**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR  
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRVFQVLLESPEFFQAEQV  
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIIISPNSYFRVLTRKR  
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY  
RVQISEDSPVGFLVVKVSATDVDGTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF  
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS  
DLDSGENGKISCSIQEDLPFLKSAENFYTLTTERPLDRESRAEYNITITVTDLGTPMLITQ  
LNMTVLIADVNDNAPAFQTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPQDP  
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS  
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGFGVWAH  
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYPPLPEAAPTQAQ  
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMMSGTRT  
LSQSYQYEVCLAGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

### **Important features:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 685-712

#### **Cadherins extracellular repeated domain signature.**

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

#### **ATP/GTP-binding site motif A (P-loop).**

amino acids 285-292

#### **N-glycosylation site.**

amino acids 418-421, 436-439, 567-570 and 786-789

## FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC  
GCCGATTGCCTCTCGGCCTGGGCAATGCTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT  
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG  
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCCCTCTCCAGGTG  
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGGCCAGGACAGGGCAGC  
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG  
TGATTCTTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT  
GGAGGAGCGGAGGACTCAAGGTGCAACGTCAGAGAGCCTTTTCTCTCTGGATGGCGCTGG  
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG  
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG  
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT  
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT  
GCCGCTTTTCTGCCAGTTTGGCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT  
CACTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC  
TGTTCCCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATAACAGATC  
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT  
GTGGTGGTAACTCAAGCCGACCAAAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA  
CTGGTTGCTTGTATTTTCTTATTCTTTTTAATTAGTTTTATTATGTATGCTACCATTTCGAA  
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT  
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA  
CATTTTCTCCAGTGACGTGTTGACTTGAACTTCAGGCAGATTAAAGAATCATTTGTTGAA  
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC  
AAAAATATTCAATAG

094553 0700  
090229 232500

## **FIGURE 165**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE  
ELLHDPMGQDRAAEFEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC  
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE  
NFTLKILNMSQDLMDFLNPNNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS  
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETCLKIFIFNQTGIEAKKNVVVTQADQ  
IGPLPSTLIKSVDWLLVFSLSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

### **Important features:**

#### **Signal peptide:**

amino acids 1-25

#### **Transmembrane domain:**

amino acids 321-340

#### **Homologous region to dilsufide isomerase**

amino acids 212-302

#### **N-glycosylation site.**

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281  
and 293-296

#### **Thioredoxin domain**

amino acids 211-227

SECRET

SECRET

**SECRET**

><MW: 33518, pI: 7.74, NX(S/T): 0

### Important features:

amino acids 1-20

amino acids 6-153

# 2025

# 2025

## **FIGURE 169**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV  
LSMEQINWLSLVYLVVSTPFGVAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN  
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPLV  
KKGEDIPMLGVYITIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA  
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK  
HFTEATKIGLCLFSLACVPFALVSQLOGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV  
GEGAATGMIFVLGQAEGILIMLAMTALTVRRSEPSLSTCQQGEDPLDWTVSLLL MAGLCTFF  
SCILAVFFHTPYRRLQAESGEPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG  
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

### **Important features:**

#### **Signal peptide:**

amino acids 1-44

#### **Transmembrane domains:**

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,  
280-300, 318-337, 341-357, 375-387, 420-441

#### **N-glycosylation site.**

amino acids 40-43 and 43-46

#### **Glycosaminoglycan attachment site.**

amino acids 468-471

[illegible]

GTCCACATCCTGCTCAACTGGGTCAAGTCCCTCTTAGACAGCAGCTCTTGTCATCATTTGCTGAAGTGGACCAAC  
TAGTTCCTCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA  
TGGCCTTGCTTGGGGTCTGCTTGTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG  
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTTGTCCAGCCTTGAAGAAGCTCTAGTGGTT  
TCTGAATCTAGCCCACCTTGGCGGTAAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTGGGGCCAGGTGG  
CTACTTATTTCTTTTAGGGGATGTGAGGAGTGACCACTCTCACGGTGAATACCAAGTGTTCAGAGGAAGTGCC  
ATCTGGTACAGTGATCGGGAAGCTGTCCAGGAATGGGGCCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCCTT  
CCAGGTGTTGCGAGCTGCCTCAGGCGCTCCCCATTAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG  
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCCTTTGATGTGCTTGCCACAGGGGATT  
GGCTCTGATCCATGTGAGATCCAAGTGCTGGACATCAATGACCACAGCCAGCGTTTCCCAAAGGCGAGCAGGA  
GCTGGAATCTCTGAGAGCGCTCTCTGCGAACCCGAGTCCCCCTGGACAGAGCTTTGACCCAGACACAGGCC  
TAACACCCCTGCACACTACATCTGTCTCCCAGTGAGCACTTTGGCTTGGATGTGATTGTGGGCCCTGATGAGAC  
CAACATGCGAAGCTCATAGTGTGAAGGAGCTGGACAGGGAAATCCATTCAATTTTTGATCTGGTGTAACTGC  
CTATGACAATGGGAACCCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG  
CCCTGCGTTTGTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT  
GACCGCCACAGACCCTGACCAGGCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCACATGCTCCAGAGGT  
GCTGGACACCTTTCAGTATTGATGCCAAGCAGGCCAGGTCATTCTGCGTCGACCTCTAGACTATGAAAAGAACCC  
TGCTTACGAGGTGGATGTTTCAGGCAAGGGCATTTGGTCCCAATCTATCCAGCCCATTTGCAAAGTTCTCATCAA  
GGCTCTGGATGTCAATGACAACATCCCAAGCATCCACGTACATGGGCCTCCAGCCATCACTGGTGTGAGAAGC  
TCTTCCCAAGGACAGTTTTTATTGCTCTTGTCTATGGCAGATGACTTGGATTGAGGACACAATGGTTTGGTCCACTG  
CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC  
CACACTGGACAGAGAGCAGTGGCCCAAAATATACCTCACTCTGTAGCCCAAGACCAAGGACTCCAGCCCTTATC  
AGCCAAGAAACAGCTCAGCACTTCAGATCAGTGACATCAACGACAATGCACCTGTGTTTGAAGAAAGCAGGTATGA  
AGTCTCCAGCGGAAAACAACCTTACCCTCTCTTCACTCAATACCATCAAGGCTCATGATGCAGACTTGGGCAT  
TAATGCGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA  
GGTCACTGCTCAGAGGTCACTGAACATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG  
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTGAGCCTCTTGGATGCCAATGATAATGCCACAGAGTGGT  
CCAGCCTGTGCTCAGCGATGGAAAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC  
CATCGAGACTCCCAATGGCTTGGGCCAGCGGGCATGACACCTCCACTGGCCACTCACAGCTCCCGGCCATT  
CCTTTTGACAACATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCCTTACAGCATCCGCAATGG  
AAATGAAGCCCACCTCTTCATCCTCAACCCTCATACGGGGCAGCTGTTTGTCAATGTACCAATGCCAGCAGCCT  
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTT  
GAGGGTCATGTTTGTACCAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCTTGGACATGTGAT  
GCTGACGGTGTATCTGCTGGCTGCTGTTGGGCATCTTCGGTTGATCTGGCTTGTGTCATGTCCATCTGCCG  
GACGAAAAGAAGGACAACAGGGCCCTACAAGTCTCGGGAGGCCAGTCCACCTACCGCCAGCAGCCCAAGAGGCC  
CCAGAAAACATTGAGAAGGCAGACATCCACCTCGTGCTGTGCTCAGGGGTGAGGCAGGTGAGCCTTGTGAAGT  
CGGGCAGTCCCACAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT  
CCACCTCACCCCGACCCTGTACAGGACGCTGCGTAATCAAGGCAACCGAGGAGCACCGGCGGAGAGCCGAGGTT  
GCTGCAAGACACGGTCAACCTCTTTTCAACCATCCAGGCAGGAATGCCTCCGGGAGAACCTGAACCTTCC  
CGAGCCCCAGCCTGCCACAGGCCAGCCAGCTTCCAGGCCTCTGAAGTTGACGGCAGCCCCACAGGGAGGCTGGC  
TGGAGACAGGGCAGTGAGGAAGCCCCACAGAGGCCACAGCCTCCTCTGCAACCTTGAACCGCAGCGACATCT  
CAATGGCAAAGTGTCCCCTGAGAAAGAATCAGGGCCCCGTGAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC  
TGCCTTCGCCAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTCAGCAAATCTCCAGCTGCT  
GTCCTTGCTGCATCAGGGCCAATTCAGGCCAAACCAACCCAGGAAATAAGTACTTGGCCAGACCGAGGAGG  
CAGCAGGAGTGCATCCGACACAGATGGCCCCAAGTGCAAGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA  
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC  
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCTCAC  
CACCACCTACCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCCG  
CAAGGCAGAGGCCACAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGTGAGCTCACT  
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCCGTGGAGGCCGCTCCGAGGCGCTGCGGCGCTCTCGGT  
CTGCGGGAGGACCTCAGTTTAGACTTGGCCACAGGTGCAGCTCAGGCATGAAAGTGAAGGGGACCCAGGTGG  
AAAGCGGGGACTGAGGGCAAGAGCAGGAGCAGCAGCAGCAGCAGCAGGTCCTGTGAACATACCTCAGACGCCT  
CTGGATCCAAGAACCAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAATCTTGTAACTCACTAGCTAG  
CGGCGGCTTGAAGCTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC  
TGACCAAAGCAGCCCCCTTGTAAAGCAGCTCTGAGTCTTTTGGAGGACAGGACGTTTGTGGCTGAGATAAGTGT  
TCCTGGCAAAAACATATGTGGAGCACAAGGGTCAGTCTCTGCGAGAAGATGCCACGGAGTATCACAGGCAGG  
AAAGGTGAGGCTTCTTGGGTAGCAGGAGTCAGGGGGCTGACCTGGGGGTGCCAGGAATGCTCTCTGACCTAT  
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 171**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA  
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFVVLATGDLALIHVEIQ  
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV  
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS  
LALAIQEDAAPGTLIIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL  
DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF  
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD  
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADALGINGKVS  
YRIQDSPVAHLVAIDSNTGEVTAQORSLNYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDA  
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTT  
IVARDADSGANGEPLYSIIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS  
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK  
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA  
GWDPCQLQAPFHLTPTLYRTLNRNQGNGAPAESREVLQDTVNLLFNHPRQRNASRENLNLP  
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ  
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGS  
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD  
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL  
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAAASGMKVQGDPPGKGTGTEGKSRGSS  
SSSRCL

### **Important features:**

#### **Signal peptide:**

amino acids 1-13

#### **Transmembrane domain:**

amino acids 719-739

#### **N-glycosylation site.**

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

#### **Cadherins extracellular repeated domain signature.**

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

## FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGG  
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG  
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT  
CCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT  
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCAATTTCTTGACCAT  
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT  
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTACATTGGC  
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACACGACTGCTTTTTTCTGTCTCTT  
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCCAA  
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGA CTCTC  
ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCT  
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA  
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA  
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG  
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC  
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC  
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT  
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA  
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT  
TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC  
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC  
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA  
TACCGCACCATAATCACTGAAGTCCTTGGAGA ACTGCAGTTCAACTTCTATCACCGTTGGTT  
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC  
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT  
GGTTTCAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC  
AAACAAAATGCTATGGTAGCATTTTTTCACCTTCATAGCATACTCCTTCCCCGTGAGGTGATA  
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGGAGAACTAACTCAAGACAATACTCA  
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAA  
AGGTGAAAAATACACTGGA ACTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT  
AGGATTTCCGTTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT  
AAAATCAGAGACTGTAACAAAAA AAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG  
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

55  
54  
53  
52  
51  
50  
49  
48  
47  
46  
45  
44  
43  
42  
41  
40  
39  
38  
37  
36  
35  
34  
33  
32  
31  
30  
29  
28  
27  
26  
25  
24  
23  
22  
21  
20  
19  
18  
17  
16  
15  
14  
13  
12  
11  
10  
9  
8  
7  
6  
5  
4  
3  
2  
1

## **FIGURE 173**

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI  
LGVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF  
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI  
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ  
QEVDAL EELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF  
DRVGKTDVPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS  
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVA  
LSSILFLYLAHKQAPEKQMAP

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Potential transmembrane domains:**

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,  
425-444

#### **N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

#### **Eukaryotic cobalamin-binding proteins**

amino acids 151-160

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

1. The first part of the document is a title page. It contains the title "The Role of the State in the Development of the Economy" and the author's name "John Doe".

2. The second part of the document is an abstract. It summarizes the main points of the paper, stating that the state plays a crucial role in the development of the economy, particularly in the areas of infrastructure, education, and healthcare.

3. The third part of the document is the introduction. It discusses the importance of the state in the development of the economy and the challenges it faces. It also mentions the author's research methodology and the scope of the study.

4. The fourth part of the document is the main body of the paper. It is divided into several sections, each discussing a different aspect of the state's role in the economy. The sections are:

- 4.1. The Role of the State in Infrastructure Development
- 4.2. The Role of the State in Education
- 4.3. The Role of the State in Healthcare
- 4.4. The Role of the State in the Environment
- 4.5. The Role of the State in the Labor Market

5. The fifth part of the document is the conclusion. It summarizes the findings of the study and discusses the implications for the future. It also mentions the author's recommendations for the state.

6. The sixth part of the document is the bibliography. It lists the sources used in the study, including books, articles, and websites.

7. The seventh part of the document is the appendix. It contains additional information related to the study, such as data tables and figures.

8. The eighth part of the document is the index. It provides a list of keywords and their corresponding page numbers, making it easier for readers to find specific information.

9. The ninth part of the document is the cover page. It contains the title, author's name, and a brief description of the document.

10. The tenth part of the document is the back cover. It contains the author's contact information and a brief description of the document.

[illegible][illegible]

[illegible]

TCCTCGCGCAGGGATCGTCCC**ATG**CGCCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCCTCCGGGAATTGCTACCTTTT  
TGGCTCCCTGCTCGTGAAGTGTCTTCTCACGGGCTGTGCCTTCAATCTGGACGTGATGGGTGCCTTGCAGCAA  
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG  
GCTGCTGGTGGGTGCTCTCCCGAGGCCCTGGCTCTTCTTGGGCAGCAGGCGAATCGCATGGAGGCCCTCTTCGCTTA  
CCCCTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACAGGGAGCTGATATGCAAAAGGAAGACAA  
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCAGTA  
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTGCCTGCTTTGTGCTCAGCCAGGA  
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAAT  
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCTTCTCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGA  
CTATAAATTGGAAGGGCAGCGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACG  
TCCCTACGAGGGCGGGGGAGAGAAGGAGCAGGACCCCGCTCATCCGGTCCCTGCCAACAGCTACTTTGGCTT  
CTCTATTGACTCGGGGAAAGTCTGGTGGCTGCAGAGAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCCAACCA  
CAAGGGTGCTGTGGTCATCCTGCGCAAGGACAGCGCCAGTCGCCTGGTGGCCGAGGTTATGCTGTCTGGGGAGCG  
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGG  
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGCTGTGTATGTGTACTTGAACCAGGGGGGTCACTG  
GGCTGGGATCTCCCCTCTCCGGCTCTGGGCTCCCCTGACTCCATGTTCCGGGATCAGCCTGGCTGCTCTGGGGGA  
CCTCAACCAAGATGGCTTTCCAGATATGTCAGTGGTGCCCTTTGATGGTGATGGGAAAGTCTTCACTACCA  
TGGGAGCAGCTGGGGGTTGTGCGCAAACCTTCAAGGTCAGAGTGTGAGGGGCAGGCTGTGGGCATCAAGAGCTTCGG  
CTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCTGACCTGCTGGTGGGCTCCCTGGCTGACAC  
CGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT  
GGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC  
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGCACACAGACCGGAGGCTCCGGGGCCAGGT  
TCCCCTGTGACGTTCTCTGAGCCGTAACCTGGAAGAACCCAGCACACGAGCCCTCGGGACCGTGTGGCTGAAGCA  
CCAGATGACCGAGTCTGTGTGAGACGCCATGTTCCAGTCCAGGAAATGTCAAAGACAAGCTTCGGGCCATTGT  
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGG  
CCCCATCCTCAATGCCCCACCAGCCAGCACCCAGCGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGA  
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCACACC  
CTGCCCCATGATGTGGATGGAAACAACAGCCCTGTTTGCACTGAGTGGGCAGCCAGTCATTGGCCCTGGAGCTGAT  
GGTCACCAACCTGCCATCGGACCCAGCCACCGCCAGCTGATGGGATGATGCCATGAAGCCAGCTCCTGGT  
CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCTGCGGAGAAGCCACTCTGCCTGTCCAA  
TGAGAATGCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCCAGGTACCTTCTACCTCAT  
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGAACTGGAGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA  
GGAGCTGCATCCAGTCTCTGACGAGGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC  
CCAGCAACTCTTCTTCTCTGGTGTGGTGGGGCCGAGAGGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT  
CAAGTATGAGGTACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTG  
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGG  
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA  
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGC  
TGAGAAGAAAGAAAAACATCACCTGGACTCGCGCCCGGGCGAGCCGCCAACCTGTGTGGTGTTCAGTGTCCCACTTA  
CAGCTTTGACCGCGCGCTGTGCTGCTGCTGCTGCGGCGGTCTCTGGAACAGCACCTTCTGGAGGAGTACTCAG  
TGTGAAGTCCCTGGAAGTGATTGTCTGGGCCAACATCAGATGAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA  
TGCTTCCACAGTGATCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT  
CATCTTCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGCTGGTGTGCTCCTGTGGAAGATGGGATTCTT  
CAAACGGCGGAAGCACCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGCA  
GTTCAAGGAGGAGGAAGACGGGCACCATCCTGAGGAACAACATGGGGCAGCCCCCGGCGGGAGGGCCCGGATGCACA  
CCCCATCCTGGCTGTGCTGACGGGCATCCCCAGCTGGGCCCCGATGGGCATCAGGGCCAGGCACCGCC**TAG**GTTC  
CATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCGAGATGCGCTCCTTGGGATGAAGAGGGTAGAGT  
GGGCTGCTGGTGTGCGATCAAGATTGGCGAGGATCGGCTTCTCAGGGGCGACAGACCTCTCCACCCACAAGAAC  
TCCTCCACCCAACTTCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG  
TGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGT  
GTAAACAGGACCCCAAGGACTAGTCTCCCCGGAAGTGCCCTTAACCTAGAGGGTCCGGGAGGAGGTTGTGTCACTGA  
TCTCAGGTGCTCCTTCTCTAGTCTTCCCCTCTCATCTGACCTTAGTTTGTCTGCCATCAGTCTAGTGGTTTCGTGGT  
TTCGCTCTATTATTATAAAAAATATTTGAGAACAAAAA

## **FIGURE 177**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL  
QPRPQSWLLVGAPQALALPGQQANRTGGFLFACPLSLEETDCYRVDIDQGADMOKESKENQWL  
GVSVRSQGPGGKIVTCAHRYEARQQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGWKFCG  
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPYEA  
GGEKEQDPRLLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV  
PEVMLSGERLTSGFGYSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVYVYLNQGGHWAGI  
SPLRLCGSPDSMFGISLAVLGD LNQDGFDPDI AVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE  
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ  
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK  
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSLSLQTPRLRRQAPGQGLPPVAP  
ILNAHQ PSTQRAEIHFLKQCGCEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA  
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRA LDPAEKPLCLSN  
ENASHVECELGNPMKRGAQVTFYLILSTSGIS IETTELEV ELLLATISEQELHPVSARARVF  
IELPLS IAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVT VSNQGGQSLRTLGS AFLNIM  
WPHEIANGKWL LYPMQVELEGGQGPQKGLCS PRPNILHLDVDSRDRRRRRELEPPEQQEPGE  
RQEPSMSWWPVSSAEKKKNITLDCARGTANC VVFS CPLYSFDRAAVLHVWGRLWNSTFLEEY  
SAVKSLEIVIRANITVKSS IKNLMLRDASTVIPVMVYLDPM AVVAEGVPWWVILLAVLAGLL  
VLALLVLLLWKMGGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGS PRREGP  
DAHPILAADGHPELGPDPGHPGPGTA

### **Important features:**

#### **Signal peptide:**

amino acids 1-33

#### **Transmembrane domain:**

amino acids 1040-1062

#### **N-glycosylation sites.**

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

#### **Integrins alpha chain proteins.**

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

## FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA  
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGTGGCT  
CACAAACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA  
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC  
TGGATGATAAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC  
AAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCTTTCGA  
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT  
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG  
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG  
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA  
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC  
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT  
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA  
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC  
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAAACTATGACCT  
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG  
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGTCTAC  
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA  
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC  
CAACACAATGTCATGGCAGTGTGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC  
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT  
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG  
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTGAC  
CATGATGTATACATTGATGATGACAGTTGAAATCAATAAATCTACATTTCTAATATTTA  
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA  
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACCTTTATGTTTAAAT  
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAGAAAATACATATGCAGTCTA  
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTACGAGAACAACCTTTGT  
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAG  
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG  
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG  
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA

## **FIGURE 179**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR  
DEVEDDYFRTWSPGKPFQALDPAKDPCCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA  
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP  
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKTKTLLRPERSRFDTSILPI  
CKDSLGMWFMRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ  
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGQCWCVDRYGNEVMGS  
RINGVADCAIDFEISGDFASGDFHEWTDDDEDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

### **Important features:**

#### **Signal peptide:**

amino acids 1-16

#### **Leucine zipper pattern.**

amino acids 246-267

#### **N-myristoylation sites.**

amino acids 357-362, 371-376 and 376-381

#### **Thyroglobulin type-1 repeat proteins**

amino acids 353-365 and 339-352

# FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC  
 CTCTTTTCAGCCCCGGGATCGCCCCAGCAGGGAATGGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC  
 GCTCTGCCCTCCGGTGCTGCTGCCTGGGGCGGCGGCTTACACCTTCCCTCGATAGCGACTTCACCTTTACCTT  
 CCCGCCGGCCAGAAGGAGTGCTTCTACAGCCCATGCCCTGAAGGCTCGCTGGAGATCGAGTACCAAGTTTTA  
 GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA  
 TCAGATGGAGTTCACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTACGACCATT  
 TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG  
 AAATATATTACTGGCACAGATATATTGGATATGAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC  
 AGACTAAGCAAAAGTGGGCACATACAAATTCTGCTTAGAGCATTGAAGCTCGTGATCGAAACATACAAGAAAGC  
 AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCAATGGTGGTGGTGTGAGCCATTCAAGTTTAT  
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACTTAAACTCCAAACTAGAGTACGTAACATTGAAA  
 AATGAGGCATAAAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATA  
 AAAGTAGGAAACAGGTATAATTTTAAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG  
 TTGTACTTAAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT  
 TTTCTTAACCTTGAAAAATTTTGCAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC  
 AGTCTGTTTTTAAACAGGTTCTATTACCCAGAACCTTTTGTAAATGCGGCAGTTACAAATTAACGTGGGAAGTTT  
 TCAGTTTTTAAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCCAA  
 CTTTTCTCTATTTACATATGCATCTCTCTATAATGTAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG  
 AGATTTTTTATAACCAATACATTTTCAAGTGAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC  
 CCAAAGCTGACATTTTTCAGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG  
 AAATGAAGAATATAGTTTAAAAGCTTCTCTCCATAGGGACACATTTTCTCTAACCTTAACTAAAGTGTAGGA  
 TTTTAAAAATTAATGTGAGGTAAAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA  
 TAATCATGTTATGTTAATTTTAAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA  
 TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA  
 AGAATTTAGAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTATAAATAAATATCTAGA  
 ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG  
 AAACCTTTGGCTGTAGGTTTTTATTTTCTACAAGAATTTCTGGTTTGAATTATTTTGTAAAGCAGGTACATTTTATA  
 AAATGTAAGCCCTACTGTAAAGGTTTAGCACTGGGTGTACATATTTATTAAAAATTTTATTATAACAACTTTTAT  
 TAAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTTAA  
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTGAAGTCTATGGGGGTCTTAC  
 TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATAATTTTAAAGTTATGCCATTTATAACGTTGTTTAT  
 GACTACATTTGTGAGTTAGAAACAAACTTAAAATTTGGGGTATAGAACCCTCAACAGGTTAGTAATGCTGGAATT  
 CTTGATGAGCAATAATGATAACCAGAGAGTGATTTTCACTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC  
 TCTTAGGCCCTGGGAGAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAATGCCGTAT  
 ATGATCAATTACCTTAATTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA  
 AGGTCAATAAGATCCTTGCCATATGAATACCCCTCCCTTTTGCCTGTAAATTTGCAATGAGAAGCAAAATTTACA  
 GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAATTTCTA  
 CCTCTCCTGTATGGCTGTTACTGTACTCTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT  
 ACATGTATGATTTTGTGCCACTGATCTTAAACCTATGATTCAAGTAACTTCTTACCATATAAAAACGATAATTGCTT  
 TATTTGGAAAAGAAATTTAGGAATACTAAGGACAATTATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG  
 CATAACCAAAAAGCAAACTTGTAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT  
 CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTT  
 AACAGCTCATTTTGTCTTTTTCAATATACAAATTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTC  
 CATAATGTAGCAGTTACCGTGTTACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT  
 GTTATGCTGTTCTTTTATGTGAATGTCAAGACATGGAGGGTGTGTTGTAATTTTATGGTAAAATTAATCCTTCTTA  
 CACATAATGGTGTCTTAAAATTGACAAAAATGAGCACTTACAATTTGTATGTCTCCTCAAATGAAGATTCTTTAT  
 GTGAAATTTTAAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG  
 CTCAAACGCTTTTACTTTATAAACAGCCATCTTAAATAAGCAACGTATTTGTGAGTACTGATATGTATATAATAA  
 AAATTATCAAAGGAAAA

500  
 400  
 300  
 200  
 100  
 0

姓名	性别	年龄	职业	住址	联系电话	备注
张三	男	45	教师	XX市XX区XX路XX号	138XXXXXXX	
李四	女	32	护士	XX市XX区XX路XX号	139XXXXXXX	
王五	男	58	工人	XX市XX区XX路XX号	137XXXXXXX	
赵六	女	28	学生	XX市XX区XX路XX号	136XXXXXXX	
孙七	男	65	退休	XX市XX区XX路XX号	135XXXXXXX	
周八	女	40	医生	XX市XX区XX路XX号	134XXXXXXX	
吴九	男	35	程序员	XX市XX区XX路XX号	133XXXXXXX	
郑十	女	25	销售	XX市XX区XX路XX号	132XXXXXXX	
冯十一	男	50	农民	XX市XX区XX路XX号	131XXXXXXX	
陈十二	女	38	公务员	XX市XX区XX路XX号	130XXXXXXX	
林十三	男	60	工程师	XX市XX区XX路XX号	129XXXXXXX	
徐十四	女	22	会计	XX市XX区XX路XX号	128XXXXXXX	
马十五	男	48	律师	XX市XX区XX路XX号	127XXXXXXX	
朱十六	女	30	记者	XX市XX区XX路XX号	126XXXXXXX	
李十七	男	55	作家	XX市XX区XX路XX号	125XXXXXXX	
王十八	女	27	设计师	XX市XX区XX路XX号	124XXXXXXX	
张十九	男	62	教授	XX市XX区XX路XX号	123XXXXXXX	
赵二十	女	33	歌手	XX市XX区XX路XX号	122XXXXXXX	
孙二十一	男	42	画家	XX市XX区XX路XX号	121XXXXXXX	
周二十二	女	29	舞蹈家	XX市XX区XX路XX号	120XXXXXXX	
吴二十三	男	53	音乐家	XX市XX区XX路XX号	119XXXXXXX	
郑二十四	女	36	演员	XX市XX区XX路XX号	118XXXXXXX	
冯二十五	男	47	导演	XX市XX区XX路XX号	117XXXXXXX	
陈二十六	女	24	编剧	XX市XX区XX路XX号	116XXXXXXX	
林二十七	男	59	制片人	XX市XX区XX路XX号	115XXXXXXX	
徐二十八	女	31	制片人	XX市XX区XX路XX号	114XXXXXXX	
马二十九	男	44	制片人	XX市XX区XX路XX号	113XXXXXXX	
朱三十	女	26	制片人	XX市XX区XX路XX号	112XXXXXXX	
李三十一	男	56	制片人	XX市XX区XX路XX号	111XXXXXXX	
王三十二	女	34	制片人	XX市XX区XX路XX号	110XXXXXXX	
张三十三	男	46	制片人	XX市XX区XX路XX号	109XXXXXXX	
赵三十四	女	28	制片人	XX市XX区XX路XX号	108XXXXXXX	
孙三十五	男	61	制片人	XX市XX区XX路XX号	107XXXXXXX	
周三十六	女	37	制片人	XX市XX区XX路XX号	106XXXXXXX	
吴三十七	男	49	制片人	XX市XX区XX路XX号	105XXXXXXX	
郑三十八	女	32	制片人	XX市XX区XX路XX号	104XXXXXXX	
冯三十九	男	51	制片人	XX市XX区XX路XX号	103XXXXXXX	
陈四十	女	25	制片人	XX市XX区XX路XX号	102XXXXXXX	
林四十一	男	63	制片人	XX市XX区XX路XX号	101XXXXXXX	
徐四十二	女	39	制片人	XX市XX区XX路XX号	100XXXXXXX	
马四十三	男	43	制片人	XX市XX区XX路XX号	099XXXXXXX	
朱四十四	女	23	制片人	XX市XX区XX路XX号	098XXXXXXX	
李四十五	男	57	制片人	XX市XX区XX路XX号	097XXXXXXX	
王四十六	女	35	制片人	XX市XX区XX路XX号	096XXXXXXX	
张三十七	男	45	制片人	XX市XX区XX路XX号	095XXXXXXX	
赵四十八	女	27	制片人	XX市XX区XX路XX号	094XXXXXXX	
孙四十九	男	60	制片人	XX市XX区XX路XX号	093XXXXXXX	
周五十	女	38	制片人	XX市XX区XX路XX号	092XXXXXXX	
吴五十一	男	50	制片人	XX市XX区XX路XX号	091XXXXXXX	
郑五十二	女	33	制片人	XX市XX区XX路XX号	090XXXXXXX	
冯五十三	男	47	制片人	XX市XX区XX路XX号	089XXXXXXX	
陈五十四	女	24	制片人	XX市XX区XX路XX号	088XXXXXXX	
林五十五	男	62	制片人	XX市XX区XX路XX号	087XXXXXXX	
徐五十六	女	40	制片人	XX市XX区XX路XX号	086XXXXXXX	
马五十七	男	44	制片人	XX市XX区XX路XX号	085XXXXXXX	
朱五十八	女	26	制片人	XX市XX区XX路XX号	084XXXXXXX	
李五十九	男	56	制片人	XX市XX区XX路XX号	083XXXXXXX	
王六十	女	34	制片人	XX市XX区XX路XX号	082XXXXXXX	
张三十一	男	46	制片人	XX市XX区XX路XX号	081XXXXXXX	
赵六十一	女	28	制片人	XX市XX区XX路XX号		

```
><subunit 1 of 1, 229 aa, 1 stop
```

MGDKIWLFPFPVLLLAALPPVLLPGAAGFTPSLSDSFTFTLTPAGQKECFYQPMPLKASLEIEY  
QVLGDAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFEL  
ILDNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDN  
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Signal peptide:

Transmembrane domain:

N-myristoylation site.

Tyrosine kinase phosphorylation site.

amino acids 55-62

116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934

CCATCCCTGAGATCTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCCAGAT  
CTCACCAGAGAGTTCGACAGACACTATGCTGCCTCCCATGGCCCTGCCAGTGTGTCTTGGATG  
CTGCTTTCCTGCCTCATTCTCCTGTGTCAAGGTGAAGAAACCCAGAAGGAACTGCC  
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT  
TTTTGTACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA  
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCTCCCTGGTGAGGAGCATTAG  
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG  
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC  
TCCACCATCTTAAACCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG  
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTTAGGGCAGGT  
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC  
TCACCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTATGATCCTCC  
TTCTTTTTCTTTTTCTTCACCTTCATTTACGGCTTTTCTCTGTCTTCCATGTCTTGAGATC  
TCAGAGAATAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

[illegible]

TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAPYVCKFKD

## amino acids 146-171

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

CCAGTCTGTTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC  
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTCTGGC  
TGCTTGC GGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCACAGGAGTGT  
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC  
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA  
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCGTGTCCTGCTGCAATACTG  
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGGCCCTCACGCTC  
CTCCCACTCTTGAGCCTCCGACTGTTAGAGTCCCCGCCCACCCCCATGGCCCTATGCGGCCCA  
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAAAAA

[illegible]

&lt;subunit 1 of 1, 125 aa, 1 stop

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP

Important features:

amino acids 1-17

amino acids 46-49

## FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC  
ACGGTTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT  
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG  
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC  
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC  
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA  
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT  
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG  
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT  
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT  
CATAGCACCTTGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA  
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA  
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT  
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC  
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGA  
GACACTAAACCAGCTATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACC  
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT  
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCCTG  
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA  
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTAAATTT  
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT  
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA  
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT  
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA  
AATAAAATTTAACATTTAAAAA

## FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA  
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRKRCMRH  
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG  
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ  
KDHHQASNSSRLHTCQRH

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **N-glycosylation site.**

amino acids 256-259

#### **Fungal Zn(2)-Cys(6) binuclear cluster domain**

amino acids 110-126

100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200

## FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA  
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTCTTCTCCTTCNG  
GAGTCCTTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG  
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC  
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT  
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA  
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT  
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG  
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT  
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT  
CATAGCACCTTGGATGGG

9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24



## **FIGURE 190**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRD LKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL  
HPDKNPNNPNAHGDFL KINRAYEVLKDEDLRKKYDKYGEKGLDNQGGQYESWNYRYDFGI  
YDDDPEIITLERREFDA AVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC  
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRS TVTELWTGNFVNS  
IQTAFAAGIGWLITFC SKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN  
TLEDRLAHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICS NLYVFQP  
SLAVFKGQGTKEYEIH HGKKILYDILAFAKESVNSHVTTLGPNFPANDKEPWLVDFFAPWC  
PPCRALLPELRRASNLL YGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHS  
AEQILEFIEDLMNP SVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART  
LTGLINVGSIDCQQYHS FCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG  
FLPQVSTDLTPTQTFSE KVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVDC  
QAYAQTCQKAGIRAYPT VKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGRNKDEL

### **Important features:**

**Endoplasmic reticulum targeting sequence.**

amino acids 744-747

**Cytochrome c family heme-binding site signature.**

amino acids 158-163

**Nt-dnaJ domain signature.**

amino acids 77-96

**N-glycosylation site.**

amino acids 484-487

## FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA  
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA  
GTCGTTGGTGAAGTTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA  
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC  
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGCCGAAA  
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT  
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA  
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA  
CATCCTAGGACATTTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG  
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCTTACCTCATCCCATAT  
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT  
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTCACCA  
AAAATCCAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA  
GATGGAATACTTACCAATAAGAAAATGATTTTTTGTTCATCGTATATCAATATCTTTCTGAG  
ACTACAGAAGTTTTCTTCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT  
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATGAATAAATAAGCTCCAGCCAGAGATG  
TATGCATGATAATGATATGAATAGTTTTCGAATCAATGCTGCAAAGCTTTATTTTCACATTTTT  
TCAGTCCTGATAATATTAAAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA  
ATTACCTGTCTTCTGTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCTT  
TCATGCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT  
TTTATTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT  
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTTCCAATCTGTAGCCATGCCACAG  
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT  
ATCTCAACCTGGACATATTTTAAGATTTCAGCATTTGAAAGATTTCCCTAGCCTCTTCTTTT  
TCATTAGCCCAAACGGTGCAACTCTATTCTGGACTTTTATTACTTGATTCTGTCTTCTGTAT  
AACTCTGAAGTCCACCAAAAAGTGGACCCTCTATATTTCTCCTCCCTTTTTATAGTCTTATAAGA  
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA  
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT  
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA  
CAATGGACCCAAGAGAAGAA

40052949455

## **FIGURE 192**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI  
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV  
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC  
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID  
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

### **Important features:**

#### **Signal peptide:**

amino acids 1-19

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-33 and 58-61

#### **Short-chain alcohol dehydrogenase family protein**

amino acids 165-202, 37-49, 112-122 and 210-219

[illegible]

CGGCGGCGGCTGCGGGCGCGAGGTTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCACAGCAG  
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGGCCCGGAGAGGGCCCGAGCCCCGGGGG  
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTTCGGTGTTTCATGATCCT  
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT  
CTAGGCCGCACACGGGGCCCGCTGCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG  
GCCGACTCCGATGTTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA  
CCTTCCCAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGAGCATGGAGGAGAGCGTGAGAG  
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG  
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG  
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG  
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG  
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA  
CGTGACACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT  
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCTTCGTGCGCGACCCC  
TTCGTGCGCCTGATCTCCGCCTTCCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG  
CAAGTTCGCCGTGCCCCATGCTGCGGCTGTACGCCAACCACACCGCCTGCCCGCCTCGGCGC  
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCTTTCGCCAACTTCATCCAGTACCTGCTGGAC  
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA  
CCCGTGCCAGATCGACTACGACTTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC  
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG  
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT  
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC  
GAGACTGAAAGCTTTCGCGTTGCTTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCC  
AGTTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC  
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

## **FIGURE 194**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSMILLIIVYWDSAGAAHFYLHTSF SRPHTGPPLPTPGPDRDRELT  
DSDVDEFLLDKFLSAGVKQSDLP RKETEQPPAPGSMEE SVRGYDWS PRDARRSPDQGRQQAER  
RSVLRGFCANSSLAFTP KERA FDDIPNSELSHLIVDDR HGAIYCYVPKVACTNWK RVMIVLS  
GSL LHRGAPYRDPLRIPREHVHNASAH LTFNKFWRRY GKL SRHLMKV LKKYTKFLFVRDPF  
VRLISAFRSKF ELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKV SFANFIQYLLDP  
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLET LDEDAQAQLLQLLQVDRQLRFPPSYRNRT  
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLF GYPK PENLLRD

### **Important features:**

#### **Signal peptide:**

amino acids 1-31

#### **N-glycosylation sites.**

amino acids 134-137, 209-212, 280-283 and 370-373

#### **TNFR/NGFR family cysteine-rich region protein**

amino acids 329-332

## FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA  
AAGAGGCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT  
ATGCCGGGAAGGTGGTGGTTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC  
GCCTTCGTGAACAGCGGGGCCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGC  
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG  
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC  
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT  
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA  
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA  
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA  
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG  
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG  
CCACTGGGCCCGCATGGGCCAGCCCGCTGAGGTGGGGCTGCGGCAGTGTTCCCTGGCCTCCGA  
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCAGAGCTGGGGTACGGGT  
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCCTGATTCTCTCATTT  
CTACTTGGGGCCCCCTTCCTAGGACTCTCCACCCCCAACTCCAACCTGTATCAGATGCAGC  
CCCCAAGCCCTTAGACTCTAAGCCAGTTAGCAAGGTGCCGGGTACCCCTGCAGGTTCCCAT  
AAAAACGATTTGCAGCC

00220533450

**THE** **NEW** **YORK** **PUBLIC** **LIBRARY**

<MW: 28317, pI: 6.00, NX(S/T): 1

amino acids 10-22, 81-91, 134-171 and 176-185

0.1  
 0.2  
 0.3  
 0.4  
 0.5  
 0.6  
 0.7  
 0.8  
 0.9  
 1.0  
 1.1  
 1.2  
 1.3  
 1.4  
 1.5  
 1.6  
 1.7  
 1.8  
 1.9  
 2.0  
 2.1  
 2.2  
 2.3  
 2.4  
 2.5  
 2.6  
 2.7  
 2.8  
 2.9  
 3.0  
 3.1  
 3.2  
 3.3  
 3.4  
 3.5  
 3.6  
 3.7  
 3.8  
 3.9  
 4.0  
 4.1  
 4.2  
 4.3  
 4.4  
 4.5  
 4.6  
 4.7  
 4.8  
 4.9  
 5.0  
 5.1  
 5.2  
 5.3  
 5.4  
 5.5  
 5.6  
 5.7  
 5.8  
 5.9  
 6.0  
 6.1  
 6.2  
 6.3  
 6.4  
 6.5  
 6.6  
 6.7  
 6.8  
 6.9  
 7.0  
 7.1  
 7.2  
 7.3  
 7.4  
 7.5  
 7.6  
 7.7  
 7.8  
 7.9  
 8.0  
 8.1  
 8.2  
 8.3  
 8.4  
 8.5  
 8.6  
 8.7  
 8.8  
 8.9  
 9.0  
 9.1  
 9.2  
 9.3  
 9.4  
 9.5  
 9.6  
 9.7  
 9.8  
 9.9  
 10.0

---

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

&lt;subunit 1 of 1, 180 aa, 1 stop

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGGQRPGLAPGPHQVPLDLVSRMKPYARMEEY  
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRI PVDLPEARCLCL  
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Signal peptide:

N-glycosylation site.

Homologous region to IL-17

amino acids 96-180.

[illegible][illegible]

## **FIGURE 200**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP  
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG  
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW  
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSEQRNRSEEAHRAEQLQDAEEEEK  
DDSNEEENKDSLVDDEEEKEDLGDEDEAEFEEDNLAAGVDEERSEANDQGPPGEDGVTRE  
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSHADKGL

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **Transmembrane domain:**

amino acids 191-211

#### **N-glycosylation site.**

amino acids 46-49

**Thioredoxin family proteins.** (homologous region to disulfide isomerase)

amino acids 56-72

#### **Flavodoxin proteins**

amino acids 173-187

## FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAACTCCGGTAAGTACCTAGCCCACATGATT  
TGA CTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC  
CAAATGCTATATCTATT CAGGGGCTCTCAAGAACAATGGAATATCATCCTGATT TAGAAAAT  
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT  
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG  
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC  
AGCCCTTGTCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATT CAGCATGTC ACT  
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG  
ACAGCTCAAATGAATTGGGATTTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT  
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT  
CTCTTCTAACTTATTT CAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG  
TATGGATT CACGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT  
GAGAAGAAGTTTCAATGTAAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA  
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG  
TTTAGAGAGCTTGGCCA ACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT  
GTATTTGTGACCTACAGGTAGGCTAGTATTATTTTCTAGTTAGTAGATCCCTAGACATGG  
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTGTAGATAGGGTCT  
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC  
CTCAGCCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTTGGTG  
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG  
TGATCTGCCCCGCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC  
CCCAAGCTTGAATTTTCTATTCTGCCATTGACTTGGCATTACCTTGGGTAAGCCATAAGCGA  
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT  
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT  
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT  
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACTCCTA  
CTTTTTCTTTATCAATTAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT  
TTTTTTTTTTTTTTTTTTTTTGTAGACAGAGTTTCGCTCTTGTTGCCAGGCTGGAGTGCAACGG  
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT  
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTTAGT  
AGAGACAGGGTTTCTCCATGTCTGGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC  
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT  
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTTAAATACA  
AAAATACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA  
ACAAGTATTAACATTTTGGGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT  
ACCATTTTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAACTATATTTCTCATG  
TGCTGATTTGTCTTATTTTTTTTCTACTTTCCCACTGGTGCTATTTTTATTTCCAATGGATA  
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAAATATGTGAAAAGAAAT  
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTTTCA  
TTCAATAAAAAAAACTCTTATCAAAAAA

## FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLIAVILGILCLVILVIAV  
VLGTMGVLSSPCPPNWIIYEKSCYLFMSLSNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ  
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL  
CSVPSYSICEKKFSM

### **Important features:**

#### **Type II transmembrane domain:**

amino acids 45-65

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 197-200

#### **N-myristoylation sites.**

amino acids 35-40 and 151-156

#### **Homologous region to LDL receptor**

amino acids 34-67 and 70-200.

# FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAGGGTCTCGCTCTGTCA  
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGTTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC  
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTGACTCCGTTCGGAGGAAAATGACTCCCCAG  
TCGCTGCTGCAGACGACACTGTTCCCTGCTGAGTCTGCTCTTCCCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG  
GAAGACTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG  
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCCTTCCCTGCAGCCCACCCTGCTTCCCGA  
TCCTTCCCTGACCCCAGGGGCTCTACCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC  
TATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG  
GCTCAGGGCCCCCGCTGTTAGCCACTTCTGTCACTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCAGTGCC  
GCCAGCTTACCTTCTCCTTCCACAGTCCCTCCCCACAGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC  
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCTTGAAGCATCCCCAGAAGGCCCTCAAGGAGGCCCTCGGCTGCCCCC  
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAACTGACCTCTGTGAGATTATGGGGGACATGGTGTCTTTCGAG  
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGG  
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCCAGAGGACGAAAGGC  
CGGAGCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCAGGACAAGAATTCC  
AGCCAAGTCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAGTAGCCAACCTCAGCGAGCCCCGTG  
GTGCTCACTTTCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTGAAGACCCACA  
TTGAGCAGCCCCGGGCATTGGAGCAGTGTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGC  
AACCCTTGACCTACTTTGCAGTGTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGC  
CTCCTCTCTACGTGGGCTGTGTCTCTGCCCCGGCCTGCCTTGTCAACATTGCCGCTACCTCTGCTCCAGG  
GTGCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTC  
CTGCTGGACACGAGCTTCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCAGCCAGTGCC  
ATCTTCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTG  
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT  
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAATATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG  
GGCGTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCACCTGGGCTCTTCAGC  
CTGGTGTCTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAA  
AAGTGTGCACATGTGCTGACACTGCTGGGCTCAGCCCTGGTCCCTTGGCCTGCCCTGGGCTTGATCTTCTTCTCC  
TTTGCTTCTGGCACCTTCCAGCTTGTGCTCCTCTACCTTTTCAGCATCATCACCTCCTTCCAAGGCTTCTCATC  
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGTGGCCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCC  
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATCTAGGCCTCCAGCCCACCTGCCATGTGATGAAG  
CAGAGATGCGGCTCGTCGCACACTGCCTGTGGCCCCCGAGCCAGGCCAGGCCAGGCCAGTGCAGCCGAGACT  
TTGGAAAGCCCAACGACCATGGAGATGGGCCGTGGCCATGGTGGACGAGACTCCCGGCTGGGCTTTTGAATTG  
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCACGGGACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTG  
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATTGCTGGG  
GGCCAGGCCTTGGATCTTGAAGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGTGGCTCCA  
GTTGCTCTGCTCTCGTGGTCAACCTGAGGGCACTCTGCATCCTCTGTCAATTTAACCCTCAGGTGGCACCCAGGG  
CGAATGGGGCCAGGGCAGACCTTACGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC  
AGCTCGCTTACCTCTGAGCCCAGGCCCTCCTCCTCAGCCCCCAGTCCCTCCCTCCATCTTCCCTGGGGTTT  
TCCTCCTCTCCAGGGCCTCCTTGCTCCTTCGTTTACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTGGGGA  
GTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCTGAGCCA  
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTG  
CTACCCCTGACCAAGCACAGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGA  
CCATGCCAGTCCCGTCTGGTTTCCATCCCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA  
GAGCCTGACACTCTCCTAAGAGGTTCTCTCAAGCCCCCAAATAGCTCCAGGCGCCCTCGGCCGCCATCATGGT  
TAATTCTGTCCAACAAACACACAGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG  
GTCACCTCCTGCAACATTGAGTCTGGTATGTGAGGCGTGCCTGAAGCAAGAACTCCTGGAGCTACAGGGACA  
GGGAGCCATCATCTCCTGCCTGGGAATCCTGGAAGACTTCTGTCAGGAGTCAAGCTTCAATCTTGACCTTGAAGAT  
GGGAAGGATGTTCTTTTACGTACCAATCTTTTGTCTTTTGTATTTAAAGAAAGTACATGTTTATTGTAGAGA  
ATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAAATCAGCTGTTGTAATCGCCTAGCAAAAAA  
AA

994333 6704

## **FIGURE 204**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSE  
EALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH  
QEESLAQGPPLLATSVTSWWSPQNI SLPSAASF TFSFHSPHTAAHNASVDMCELKRD LQLL  
SQFLKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD  
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLLVDFSSQALFQDNSSQVLGE  
KVLGIVVQNTK VANLT E P V V L T F Q H Q L Q P K N V T L Q C V F W V E D P T L S S P G H W S S A G C E T V R R E  
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSA LACLV TIAAYLC SRVPLPC  
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLE  
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPI ILAVHRTPEGVIY  
PSMCWIRDSLVS YITNLGLFSLVFLFNMA MLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG  
LPWALIFFSFASGTFQLVVLYLFSIITSFQGF LIFIWYWSMRLQARGGPSPLKSNSDSARLP  
ISSGSTSSSRI

### **Important features:**

#### **Signal peptide:**

amino acids 1-25

#### **Putative transmembrane domains:**

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590  
and 634-657

#### **Microbodies C-terminal targeting signal.**

amino acids 691-693

#### **cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 198-201 and 370-373

#### **N-glycosylation sites.**

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327  
and 341-344

#### **G-protein coupled receptors family 2 proteins**

amino acids 475-504

[illegible]

姓名	性别	年龄	职业	住址	联系电话	备注
张三	男	45	教师	北京市朝阳区	13800138000	
李四	女	35	医生	北京市海淀区	13900139000	
王五	男	55	工程师	上海市浦东新区	13600136000	
赵六	女	25	学生	广东省广州市	13500135000	
孙七	男	65	退休	北京市西城区	13400134000	
周八	女	40	会计	浙江省杭州市	13300133000	
吴九	男	30	程序员	江苏省南京市	13200132000	
郑十	女	50	护士	山东省济南市	13100131000	
陈十一	男	20	实习生	河南省郑州市	13000130000	
林十二	女	60	退休	四川省成都市	12900129000	
徐十三	男	48	经理	湖北省武汉市	12800128000	
冯十四	女	38	设计师	湖南省长沙市	12700127000	
朱十五	男	58	教授	安徽省合肥市	12600126000	
马十六	女	28	销售	江西省南昌市	12500125000	
宋十七	男	68	退休	福建省厦门市	12400124000	
李十八	女	42	文员	广东省深圳市	12300123000	
王十九	男	32	司机	广西壮族自治区南宁市	12200122000	
张二十	女	52	厨师	贵州省贵阳市	12100121000	
陈二十一	男	22	实习生	云南省昆明市	12000120000	
林二十二	女	72	退休	陕西省西安市	11900119000	
徐二十三	男	47	工程师	甘肃省兰州市	11800118000	
冯二十四	女	37	设计师	青海省西宁市	11700117000	
朱二十五	男	57	教授	宁夏回族自治区银川市	11600116000	
马二十六	女	27	销售	新疆维吾尔自治区乌鲁木齐市	11500115000	
宋二十七	男	67	退休	内蒙古自治区呼和浩特市	11400114000	
李二十八	女	43	文员	辽宁省沈阳市	11300113000	
王二十九	男	33	司机	吉林省长春市	11200112000	
张三十	女	53	厨师	黑龙江省哈尔滨市	11100111000	
陈三十一	男	23	实习生	内蒙古自治区包头市	11000110000	
林三十二	女	73	退休	辽宁省大连市	10900109000	
徐三十三	男	49	工程师	吉林省吉林市	10800108000	
冯三十四	女	39	设计师	黑龙江省齐齐哈尔市	10700107000	
朱三十五	男	59	教授	内蒙古自治区鄂尔多斯市	10600106000	
马三十六	女	29	销售	辽宁省锦州市	10500105000	
宋三十七	男	69	退休	吉林省延边朝鲜族自治州	10400104000	
李三十八	女	44	文员	黑龙江省牡丹江市	10300103000	
王三十九	男	34	司机	内蒙古自治区呼伦贝尔市	10200102000	
张四十	女	54	厨师	辽宁省营口市	10100101000	
陈四十一	男	24	实习生	吉林省松原市	10000100000	
林四十二	女	74	退休	黑龙江省佳木斯市	09900099000	
徐四十三	男	50	工程师	内蒙古自治区通辽市	09800098000	
冯四十四	女	40	设计师	辽宁省阜新市	09700097000	
朱四十五	男	60	教授	吉林省白山市	09600096000	
马四十六	女	30	销售	黑龙江省大庆市	09500095000	
宋四十七	男	70	退休	内蒙古自治区赤峰市	09400094000	
李四十八	女	45	文员	辽宁省盘锦市	09300093000	
王四十九	男	35	司机	吉林省辽源市	09200092000	
张五十	女	55	厨师	黑龙江省伊春市	09100091000	
陈五十一	男	25	实习生	内蒙古自治区锡林郭勒盟	09000090000	
林五十二	女	75	退休	辽宁省丹东市	08900089000	
徐五十三	男	51	工程师	吉林省四平市	08800088000	
冯五十四	女	41	设计师	黑龙江省鸡西市	08700087000	
朱五十五	男	61	教授	内蒙古自治区兴安盟	08600086000	
马五十六	女	31	销售	辽宁省抚顺市	08500085000	
宋五十七	男	71	退休	吉林省铁岭市	08400084000	
李五十八	女	46	文员	黑龙江省鹤岗市	08300083000	
王五十九	男	36	司机	内蒙古自治区乌兰察布市	08200082000	
张六十	女	56	厨师	辽宁省铁岭市	08100081000	
陈六十一	男	26	实习生	吉林省白城市	08000080000	

姓名	性别	年龄	职业	住址	联系电话	备注
张三	男	45	教师	北京市海淀区中关村大街100号	13910123456	
李四	女	32	医生	北京市朝阳区建国路123号	13801012345	
王五	男	58	工程师	上海市浦东新区世纪大道100号	13621012345	
赵六	女	28	程序员	广州市天河区珠江新城100号	13530123456	
孙七	男	65	退休	北京市西城区德胜门内大街100号	13440123456	
周八	女	40	会计	深圳市南山区科技园100号	13350123456	
吴九	男	35	销售	武汉市武昌区中南路100号	13260123456	
郑十	女	50	律师	上海市黄浦区南京东路100号	13170123456	
冯十一	男	25	学生	北京市海淀区中关村大街100号	13080123456	
陈十二	女	38	护士	北京市朝阳区建国路123号	12990123456	
林十三	男	60	工程师	上海市浦东新区世纪大道100号	12800123456	
徐十四	女	30	程序员	广州市天河区珠江新城100号	12710123456	
马十五	男	55	退休	北京市西城区德胜门内大街100号	12620123456	
朱十六	女	42	会计	深圳市南山区科技园100号	12530123456	
李十七	男	33	销售	武汉市武昌区中南路100号	12440123456	
王十八	女	52	律师	上海市黄浦区南京东路100号	12350123456	
张十九	男	27	学生	北京市海淀区中关村大街100号	12260123456	
赵二十	女	37	护士	北京市朝阳区建国路123号	12170123456	
孙二十一	男	57	工程师	上海市浦东新区世纪大道100号	12080123456	
周二十二	女	29	程序员	广州市天河区珠江新城100号	11990123456	
吴二十三	男	59	退休	北京市西城区德胜门内大街100号	11800123456	
郑二十四	女	41	会计	深圳市南山区科技园100号	11710123456	
冯二十五	男	34	销售	武汉市武昌区中南路100号	11620123456	
陈二十六	女	51	律师	上海市黄浦区南京东路100号	11530123456	
林二十七	男	26	学生	北京市海淀区中关村大街100号	11440123456	
徐二十八	女	36	护士	北京市朝阳区建国路123号	11350123456	
马二十九	男	56	工程师	上海市浦东新区世纪大道100号	11260123456	
朱三十	女	31	程序员	广州市天河区珠江新城100号	11170123456	
李三十一	男	61	退休	北京市西城区德胜门内大街100号	11080123456	
王三十二	女	43	会计	深圳市南山区科技园100号	10990123456	
张三十三	男	34	销售	武汉市武昌区中南路100号	10800123456	
李三十四	女	53	律师	上海市黄浦区南京东路100号	10710123456	
王三十五	男	28	学生	北京市海淀区中关村大街100号	10620123456	
张三十六	女	38	护士	北京市朝阳区建国路123号	10530123456	
李三十七	男	58	工程师	上海市浦东新区世纪大道100号	10440123456	
王三十八	女	29	程序员	广州市天河区珠江新城100号	10350123456	
张三十九	男	59	退休	北京市西城区德胜门内大街100号	10260123456	
李四十	女	40	会计	深圳市南山区科技园100号	10170123456	
王四十一	男	35	销售	武汉市武昌区中南路100号	10080123456	
张三十二	女	55	律师	上海市黄浦区南京东路100号	09990123456	
李四十三	男	25	学生	北京市海淀区中关村大街100号	09800123456	
王四十四	女	35	护士	北京市朝阳区建国路123号	09710123456	
张三十五	男	55	工程师	上海市浦东新区世纪大道100号	09620123456	
李四十五	女	25	程序员	广州市天河区珠江新城100号	09530123456	
王四十六	男	55	退休	北京市西城区德胜门内大街100号	09440123456	
张三十六	女	45	会计	深圳市南山区科技园100号	09350123456	
李四十六	男	35	销售	武汉市武昌区中南路100号	09260123456	
王四十七	女	55	律师	上海市黄浦区南京东路100号	09170123456	
张三十七	男	25	学生	北京市海淀区中关村大街100号	09080123456	
李四十七	女	35	护士	北京市朝阳区建国路123号	08990123456	
王四十八	男	55	工程师	上海市浦东新区世纪大道100号	08800123456	
张三十八	女	25	程序员	广州市天河区珠江新城100号	08710123456	
李四十八	男	55	退休	北京市西城区德胜门内大街100号	08620123456	
王四十九	女	45	会计	深圳市南山区科技园100号	08530123456	
张三十九	男	35	销售	武汉市武昌区中南路100号	0844	

**FIGURE 207**

MSLFGLLLLTALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR  
FPHTYPRNTVLVWRLVAVEENVWIIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC  
GSGTVPQKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL  
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKRSRVVDLNLLEEVRLY  
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ  
LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

[illegible]

# FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGAAGA  
 AGACTAAAAATCGTGTTCCTAATGTGGACACTGAAGAGACAAATCTTATCCTTTTTAACATAATCCTAATTTCC  
 AAACCTCCTTGGGGCTAGATGGTTTTCTAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG  
 ATCGTGGACTGCACAGACAAGCATTGTGACAGAAATTCCTGGAGGTATTTCCACGAACACCACGAACCTCACCTC  
 ACCATTAACCACATACCAGACATCTCCCCAGCGTCCTTTACAGACTGGACCATCTGGTAGAGATCGATTTTCAGA  
 TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC  
 TTTAGTGGACTCACTTATTTAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCG  
 CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC  
 AACATAGAAATCTCTACCTGGGCCAAACTGTATTATCGAAATCCTTGTATGTTTCATATCAATAGAGAAA  
 GATGCTTCTCTAAACTTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGCCGTCCCTACTGTT  
 TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC  
 CTCAACCAATTACAAATTCTTGACCTAAGTGGAAATTGCCCTCGTTGTTATAATGCCCCATTTCTTGTGCGCCG  
 TGTAATAATAATTCTCCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAAGTTTTACGTCTA  
 CACAGTAACCTCTTTCAGCATGTGCCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACCTGGATCTGTCC  
 CAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG  
 TCTTTCAATTTTGAACCTTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG  
 AAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA  
 AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAGA  
 CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT  
 GCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCTTGAACAATTACATTATTTAGATATGATAAGTATGCA  
 AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTTATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC  
 TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCCTGTGATTTTCAGCATCTTTCTTCTCAAATGCCTG  
 AATCTGTGAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTTG  
 GACTTCTCCAACAACCGGCTTGATTTACTCCATTCACAGCATTTTGAAGAGCTTCACAACTGGAAGTTCTGGAT  
 ATAAGCAGTAATAGCCATTATTTTCAATCAGAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGT  
 CTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA  
 ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG  
 AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGT  
 ATGCCTCCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT  
 CTAAAGAACCTGGAACCTTTGGACCTCAGCCACAACCACTGACCACTGTCCCTGAGAGATTATCCAACCTGTTCC  
 AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTCTACAAGATGCCTTC  
 CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCAGAAAATGTCTC  
 AACAACTGGAAGATGTTGCTTTTGCATCATAATCGGTTTTCTGTGCACCTGTGATGCTGTGTGGTTTTGTCTGGTGG  
 GTTAACCATACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC  
 CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA  
 TCTGTATCTCTCTTCTCATGGTGTGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT  
 TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTGCTATGATGCTTTTATTGTGTAT  
 GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAA  
 CATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA  
 CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAGACTGAAAATTTTAAGATAGCATTTTAC  
 TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG  
 TCCAAGTTCTCCAGCTCCGAAAAGGCTCTGTGGGAGTTCTGTCTTGTGAGTGGCCAAACAAACCCGCAAGCTCAC  
 CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA  
 ACGGTCTAGCCCTTCTTGTCAAAACACAACCTGCCTAGTTTACCAAGGAGAGGCCTGGC

40523' 5235

## FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG  
IPTNTTNLTTLTINHIPDISPASFHRLDHLVEIDFRNCNCVPIPLGSKNNMCIKRLQIKPRSFS  
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLQNCYYR  
NPCYVSYSEIEKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL  
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPFRWF  
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL  
KILRIRGYVFKELKSFNLSPLHNLQNLVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS  
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ  
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH  
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLOKLMMNDNDISSSTSRTMESES  
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSL  
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSCLKNLILKNNQIRSLTKYFLQ  
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHNRFLCTCDAVWFVWVWNHTEVTIP  
YLATDVTVCVGPAGHKGQSVISLDLYTCELDLTNLILFSLISISVSLFLMVMMTASHLYFWDVW  
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE  
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLE  
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHYPFWQCLKNALATDNHVAYSQVFKETV

### Signal sequence:

amino acids 1-26

### Transmembrane domain:

amino acids 840-860

# FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAAACAGAAACATGGAACATGTTCCCTTC  
 AGTCGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA  
 GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAG  
 TTCCCCAAACGGTGGGCAAAATATGTGACAGAACTAGACCTGTCTGATAATTCATCACACACATAACGAATGAAT  
 CATTTCAAGGGCTGCAAAATCTCACTAAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATC  
 CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCTCAACCTAAAAAACCTAAGGGAGTTACTGC  
 TTGAAGACAACCAGTTACCCCCAAATACCCTCTGGTTTGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAACA  
 ATATATACAACATAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAAGTCT  
 ATTTTAAACAAAGTTTTCGAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT  
 CACTATCTTTCAATTCTCTTTTACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACTTTTCTGAGCAACA  
 CCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAAC  
 GTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGATGGTGGTGCCTTCAATTAATATAGATCGTTTTG  
 CTTTTCAAACCTTGACCAACTTCGATACCTAAACCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT  
 TTAAAAATATGCTCATCTGAAGGTGCTGGATCTTGAATTTCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT  
 TTTTAAACGATGCTGCCCCGCTTAGAAATACTTGACTTGTCTTTAACTATATAAAGGGAGTTATCCACAGCATA  
 TTAATATTTCCAGAACTTCTCTAACTTTTGTCTCTACGGGCATTGCATTTAAGAGGTTATGTGTTCAGGAAC  
 TCAGAGAAGATGATTTCCAGCCCCCTGATGCAGCTTCCAAACTTATCGACTATCAACTTGGGTATTAATTTTATTA  
 AGCAAATCGATTTCAAACCTTTTCCAAATTTCTCCAATCTGGAAATTATTTACTTGTGAGAAAAACAGAAATATCAC  
 CGTTGGTAAAAGATACCCGGCAGAGTTATGCAAATAGTTTCTCTTTTCAACGTCATATCCGAAACGACGCTCAA  
 CAGATTTTGAGTTTGACCCACATTGCAACTTTTATCATTTTCAACCGTCTTTAATAAAGCCACAATGTGCTGCTT  
 ATGGAAGCCCTTAGATTAAAGCCTCAACAGTATTTTCTTCATTGGGCCAAACCAATTTGAAAACTTCTCTGACA  
 TTGCTGTGTTAAATCTGTCTGCAATAGCAATGCTCAAGTGTTAAGTGGAACGAAATTTTTCAGCCATTCTCTCATG  
 TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGCTAGTGTCTTACTGAATTGTCCGACTTGG  
 AAGTTCTAGATCTCAGCTATAATTACACTATTTTCAAGATAGCAGGCGTAACACATCATCTAGAATTTATTCAA  
 ATTTTCAAAATCTAAAAGTTTTAACTTGAGCCACAACAACATTTATACTTTAACAGATAAGTATAACCTGGAAA  
 GCAAGTCCCTGGTAGAATTAGTTTTTCAAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA  
 TCTCCATTTTCAAAGGCTCAAGAATCTGACAGCTCTGGATTATCCCTTAATAGGCTGAAGCACATCCCAATG  
 AAGCATTCCTTAATTTGCGAGCGAGTCTCACTGAAGTACATATAAATGATAATATGTTAAAGTTTTTTAACTGGA  
 CATTACTCCAGCAGTTTCTCGTCTCGAGTTGCTTGACTTACGTGGAACAAACTACTCTTTTTAACTGATAGCC  
 TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCACCTACCCTCTGGCTTTC  
 TTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTG  
 AAATAGACCAACCAAAATATCTATGTTGGAATACACGGAAACCCCTTTGAATGCACCTGTGACATTGGAG  
 ATTTCCGAAGATGGATGGATGAACATCTGAATGCTCAAAATTTCCAGACTGGTAGATGTCAATTTGTGCGAGTCTG  
 GGGATCAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTTTCAGATGTCACTGCAGTGATATTAT  
 TTTTCTTCACTGTTCTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACCATTTGTTTTACTGGGATGTTT  
 GGTATATATATAATGTGTGTTTAGCTAAGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACTTTCTATGATG  
 CTTACATTTCTTATGACACCAAGATGCCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAG  
 AGAGCCGAGACAAAAACGTTCTCCTTTGTCTAGAGGAGAGGATTGGGACCCGGGATTGGCCATCATCGAACCC  
 TCATGCAGAGCATCAACCAAGCAAGAAAAACGATTTTGTTTTAAACCAAAAAATATGCAAAAAAGCTGGAACTTTA  
 AAACAGCTTTTTACTTTGGCTTTGCGAGAGGCTAATGGATGAGAACATGGATGTGATTTATTTATCCTGCTGGAGC  
 CAGTGTACAGCATTTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACA  
 ACCCGAAGGCAGAAGGCTTGTGTTTGGCAAACTCTGAGAAATGTGGTCTTGACTGAAATGATTCACGGTATAACA  
 ATATGTATGTGATTCCATTAAGCAATACTAACTGACGTTAAGTCATGATTTTCGCGCCATAATAAAGATGCAAG  
 GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTTATCCCAAACTTAGTGGTTTAAAAACAACACA  
 TTTGCTGGCCACAGTTTTTGGGGTCCAGGAGTCCAGGCCAGCATAAAGTGGTCTCTGCTCAGGGTGTCTCAG  
 AGGCTGCAATGTAGGTGTTTACCAGAGACATAGGCATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCA  
 ATTCCTCCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCACAGGCAGCTTGCTTC  
 ATCAGAGCTAGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAAATCGAATCAAAAAAGTGAT  
 ATCTCATCACTTTGGCCATATTTCTATTTGTTAGAAGTAAACACAGGTCCCACCAGCTCCATGGGAGTGACCACC  
 TCAGTCCAGGGAAAAACAGCTGAAGACCAAGATGGTGAAGTCTGATTGCTTCAGTTGGTCATCAACTATTTTCCCT  
 TGACTGCTGTCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC  
 ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAATTTTGCCACTGTGACTAATGGTCCTAATA  
 TTAAGCTGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTCCGTTTAT  
 TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG  
 ATAGCTTTTAAAGCATCTTTTACTTCTTACCAATTTTAAAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATA  
 TTGTTAATTGCCATTGCTGTAAATCTTAAATGAATGAATAAAAAATGTTTCATTTTACAAAAA

0573535 0573535

## FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPTVG  
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFNLN  
KNLRELLLEDNQLPQIPSGLPESLTSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV  
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL  
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTLQLRYLNLSSTSLRKINAAWFKNM  
PHLKVLDFEYNYLVGEIVSGAFLTMLPRLEILDLSFNLIKGSYPQHINISRNFSKLLSLRAL  
HLRGYVFQELREDDFQPLMQPLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSENRISPLV  
KDTRQSYANSSSFQRHIRKRRSTDFFEDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI  
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDELEV  
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL  
DILWNDDDNRYISIFKGLKNLTRLDLNRLKHIPNEAFLNLPASLTEHINDNMLKFFNWT  
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDL  
SNLLKTINKSALETKT'TTKLSMLELHGPNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP  
GDQRGKSIVSLELTTCVSDVTAVILFFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK  
GYRSLSTSQTIFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPLAID  
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIIFILLEPVLQHSQYLRL  
RQRICKSSILQWPDNPKAEGFWQTLRNVLVTENDSRYNMYVDSIKQY

### Signal sequence:

amino acids 1-26

### Transmembrane domain:

amino acids 826-848

## FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT  
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA  
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC  
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC  
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA  
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA  
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT  
CTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT  
CCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCCTCA  
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC  
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAGGAC  
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA  
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA  
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG  
CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC  
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA  
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGCTGGC  
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC  
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCCTTCCTG  
GAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGAACTGCCCAGCGCCCCAGGCTG  
GACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG  
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC  
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC  
CCCTGGCTACCCCCACCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG  
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG  
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAATGAAACGTGA  
AAAGGGCGGCCGCGACTCTAGAGT  
CGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTACAAAT

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT

FIGURE 213

MRGSQEVLLMWLLVLA VVGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC  
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGA AICQPPCRNGGSCVQPGRCR  
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSL SADGTL CVPKGGPPRVA  
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG  
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

[illegible]

## FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG  
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC  
AGCAGCATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGA  
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC  
AGTGGATGAGCAACCCAACGGGGGGCCGGGGAGGGGAAGTGGCCCCGAGGGAGAGGAACCCC  
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG  
CACAGGTGGCCCCCACCACCCGAGGAGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTC  
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGC  
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA  
CCGGCCCCGGCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTG  
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC  
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA  
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT  
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA  
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG  
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCC  
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG  
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT  
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC  
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCAGCAGCTCGGCCGCATCGAC  
TCCCTGAGCGAGCAGATTTCTTCTTCTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA  
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATG  
CCCCTGCCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC  
AGGGCCTTCCTCCTCTTCTCCTCCCCTTCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT  
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA  
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC  
CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC  
CCCCAGCACATAAAAAATGAAACGTG

**FIGURE 215**

MRGSQEVLLMWLLVLA VGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTC DGHRA C  
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR  
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSL SADGTL CVPKGGPPRVA  
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG  
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

[illegible]

**THE UNIVERSITY OF CHICAGO**

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCCA  
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG  
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT  
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC  
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCG  
CTGGAGGCACAGGCC**ATG**AGGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGTGC  
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG  
GGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGAC  
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG  
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC  
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG  
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA  
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT  
GCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG  
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG  
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA  
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCC  
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCT  
GGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGGCTGGACTGAGCCCC  
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG  
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCCTCTTCTCCTCCCCTTCTCGGGAG  
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC  
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC  
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG  
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC  
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR  
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSadgtLCVpKGgPPRVA  
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG  
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

[illegible]

## FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCCTTGTGAGGAGAGACAGCCTCCCGGCCCGGGGAG  
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCTCGCCGTCAGCTGCCGGCCG  
AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCTGTTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG  
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT  
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA  
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAACTCCAGGCAGGAACACCTTTTGCTGTG  
ATGGCCGCGTCATGATGGCCCCGCAAAAGGGCATTTTCTACCTGACCCTTTTCTCATCCTGGGGACATGTACAC  
TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTGAGCTGTCTCTGCCATCCCTGTATTTGCTGCCATGCTCT  
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCTCGGGCGCTACCAGATG  
AAGCAGCTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA  
TCAAGAAATTTCCAGATAAACAACAGATTGTGAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC  
GGCCTCCCATTTGCAGCATCTGTGACAACCTGTGTGGAGCGCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTG  
TTGGAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCT  
TCAACATCGTCTATGTGCCCTCAAATCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAACTCCTGGAACCTG  
TTCTAGAAGTCTCATTTGCTTCTTTACACTCTGGTCCGTCTGGGACTGACTGGATTTTCATACTTTCTCGTGG  
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC  
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA  
TTTTGCCACTGGAGGAAAGTGGAAGTCGACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC  
CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG  
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCCTATCTATGGAAGAGACTTTTGTTTGTGTT  
TAATTAGGGCTATGAGAGATTTAGGTTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT  
GTTTTTCTTTGGTCTTTAGTCAACCCAGTTGCACACTGGCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACT  
CAAGGCAGTGGCAGAAGATGTGAGTCACTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT  
CCATGGCCTCAGCCACAGGGTCCCCCTTGAGCCCCCTCTCTTCCCTCCAGATCCCAGCCCTCCTGCTTGGGGTCAC  
TGGTCTCATTTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCAAGCTGCTGCACGTGCTGAGTCCAGA  
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTTCAGGACTGAAGAGGAGGGAGAG  
TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCACAAATCCTTTTAGGAATGGGACAGGTACCT  
TCCACTTGTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTCTTTTGACTCCTGCTCCCATTAGGAG  
CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAAC  
ACTATCCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG  
GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAACCT  
CGGCTGAGTGAGGGAAGCCAGCACTGCTGCCCTCTCGGGTAACTCACCCCTAAGGCCTCGGCCCCACCTCTGGCT  
ATGGTAACCACACTGGGGGCTTCCCTCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCAGAGCCACTT  
CACCTTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTTCAGGGAAGAAGATTTATGT  
ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT  
GCGGTGGGGGAGTGTAACCGGAACTTTTCTATCTATTTGAAGCGGATTAACTGTGTCTAATGCA

59455 59555 59655

[illegible]

•

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT  
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT  
CCCACAGAGCNCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA  
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA  
ACATCGT

[illegible]

## FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGACACAAGCTTGAGAGCAACACAA  
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA  
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC  
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC  
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT  
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA  
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG  
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC  
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATTTT  
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC  
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC  
GAATACTTGGAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC  
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCGTGAACATCCACCATACA  
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC  
TCAGCAGTCCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA  
GAAAGGGGTGAAAGTGGAACAGACCTTTCCTCTCAAACTCATCTTCTTCAATGTCTCTG  
AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC  
ATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGG  
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTTGATGTGAGTGCC  
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC  
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA  
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA  
AATTGCCTTGACAGATATTTAGGTACAATGGAGTTTTCTTTCCCAAACGGGAAGAACACAGC  
ACACCCGGCTTGGAACCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA  
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA  
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG  
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA  
AAAAA

1000 900 800 700 600 500 400 300 200 100

**SECRET**

Signal peptide:

amino acids 1-28

## FIGURE 223

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC  
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT  
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA  
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT  
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAAACGCAGTACAGCAT  
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA  
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATT  
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG  
ACCAGAG

100520 555160

[illegible]

ATGGCTGGTGAACGGCGGGGCGGGCAGGGGACCGGGCCCGGGCCCGGAGCGGGCCAGCTGCCGGGAGCCCTGA  
ATCACCGCTGGCCCGACTCCACCATGAACGTGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG  
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT  
GCACTGCTTCTGGGCTGCCTTGTGGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGCCTTACA  
GAGGCTTCGATTCGAGTGGCTGGAAAAATCTCGGATCCCTGGACCGAGGGGTGAGCCCCGTGTGAGGACTTTTAC  
CAGTTCCTCTGTGGGGCTGGATTTCGAGGAACCCCTGCCGACGTGGGCTTCTCGTGGAAACACCTTCAACAG  
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG  
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCAATTGAGGAGCTGGGAGCCCAGCCACTGAGA  
GACCTCATTGAGAAGATTGGTGGTTGGAAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAG  
GCAGTAGCAGGGACCTACAGGGCCACCCCATCTTCCACGCTCTACATCAGTGCCGACTCTAAGAGTCTTCAACAGC  
AATGTTATCCAGGTGGACCACTGTGGGCTCTTTCTGCCCTCTCGGATTACTACTTAAACAGAAGTGCCAATGAG  
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGGATGCTGTCTGGTGGGCGGGCCACCTCCACGAGG  
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGAC  
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC  
CTGTCTTTCTTGCTGTCAACATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG  
CAGGTGTACAGCTCATCAACCGCAGGAAACCAAGCATCTGAAACAATTACCTGATCTGGAACCTGGTGCAAAAG  
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGTCTGGAGACCTCTATGGCACTAAGAAG  
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC  
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG  
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT  
GATATGATTGGTTTCCCAGACTTTATCCTGGAGGCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT  
GAAGATTCTTTCTTCCAACAACTGTTGAATTTGTACAACCTTCTCTGCCAAGGTATTTGCTGGCAGCTCCGCAAC  
CCTCCAGCCGAGACAGGTGAGCATGACCCCCCAGCAGTGAATGCCCTACTACCTTCCAACCTAAGAATGAGATC  
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATC  
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG  
CGGCCCTGGTGGCAGAATGATCCCTGGCAGCCTTCCGGAACCACACGCCTGCATGGAGGAACAGTACAATCAA  
TACCAGTCAATGGGGAGAGGCTCAACGGCCGCCACGACGCTGGGGGAGAACTTACTGACAACCGGGGCTGAAG  
GCTTCCTACATGCTTACAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC  
AACCACCGACTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGTCTCGTCCGCACACCAGAGAGCTCTCACGAGGGG  
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCTGCGG  
CACTTCGCTGCCCTGTGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTTAGACCTGGATCAGGGGA  
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTGGGAGGAAGCAA  
ATGCAGAAGCTGGGCTGGGTCTAGTCCCTCCCCCAGCAGGTGACATGAGTACAGACCTCTCAATCACCACATTG  
TGCTCTGCTTTGGGGGTGCCCTGCCTCCAGCAGAGCCGCCACCATCTACCTGTGACATCTTCCGTGTCAACCT  
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

## FIGURE 225

MNVALQELGAGSNVGFQKQTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH  
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA  
ILKHLLNTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLEKIGGWNITGPWDQDN  
FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGFLPSRDYVLNRTANEKVLTA  
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP  
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL  
DRRFESAQEKLETLTYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI  
SEIRTA FEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF  
FQNMLNLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH  
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV  
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS  
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGLCEVW

**Type II Transmembrane domain:**

amino acids 32-57

**FIGURE 226**

GCCCCGCCCTCCGCCCTCCGACTCCCGCCTCCCTCCCTCCGCCGCTCCCGCGCCCTCCTCCCTCCCTCCTCCC  
CAGCTGTCCCGTTCGCGTCATGCCGAGCCTCCCGCCCCCGCGGCCCGCGCTGCTGCTCCTCGGGCTGCTGCTGCT  
CGGCTCCCGGCCGCCCGCGCGCGCGGCCAGAGCCCCCGTGTGCCCCATCCGTTCCTGAGAAGGAGCCGCTGCC  
CGTTCGGGGAGCGGCAGGTAGGTGGGCGCCCCGGGGGAGGCGCGGGCGGGGAGTCGGGCTCGGGGCGAGTCAGCGC  
CAGCCCCGAGGGGGCGCGGGGCGCAGGTGGCTCGGCGCGGGCGGGCGGCCCGAGGGTGGGCGGGGCGAGAAGGGC  
GCGGTGCCTGGGACCCGGGACCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGAGCTGGGCAGCGGCCTCCAGC  
CAAGCCCGTCCCCGCGAGGCTGCACCTTCGGCGGGGAAGGTCTATGCCTTGGACGAGACGTGGCACCCGGACCTAGG  
GGAGCCATTCCGGGTGATGCGCTGCGTGTGTGCGCTGCGAGGCGCAGTGGGGTCCGCGTACCAGGGGGCCCTGG  
CAGGGTCAGCTGCAAGAACATCAAACCAGAGTGGCCAAACCCCGGCCTGTGGGCAGCCGCGCCAGCTGCCGGGACA  
CTGTGCCAGACCTGCCCCCAGGACTTCGTGGCGCTGCTGACAGGGCCGAGGTGCGAGGCGGTGGCACGAGCCCG  
AGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCCTACAGGCGGCTGGACCGCCCTACCAGGATCCGCTT  
CTCAGACTCCAATGGCAGTGTCTGTGAGCACCCTGCAGCCCCCACCAGATGGCTGGTCTGTGGGGTGTG  
GCGGGCAGTGCCTCGGTTGTCTCTGCGGCTCCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCA  
CCCTTCAGGGGAGGTCTGGGGGCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCATCCTGAC  
TCTAGAAGGCCCCCCACAGCAGGGCGTAGGGGGCATCACCTGCTCACTCTCAGTGACACAGAGGACTCCTTGCA  
TTTTTTGCTGCTCTTCCGAGGCTTGCAGGACTTAACCCAGGTTCCCTTGAGGCTCCAGATTCTACACCAGGGGCA  
GCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCAGGAACCAGGCTTTGCTGAGGTGCTGCCAACCTGACAGT  
CCAGGAGATGGACTGGCTGGTGTGTTGGGGAGCTGCAGATGGCCCTGGAGTGGGCAGGCAGGCCAGGGCTGCGCAT  
CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGCTCTGCAAAGTGTCTTTGTGGGGCTAATGCCCTGATCCC  
AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCTTGATCCTCCAGGTGCAATT  
GGTAGGGACAACAGTAGGTGGTGGCCATGACACTGGAAACCAAGCCTCAGCGGAGGGATCAGCCCACTGTCTCT  
GTGCCACATGGCTGGCCTATCCTCCCTGCCCGCAGGCCGTGGGTATCTGCCCTGGGCTGGGGTGCCCGAGGGGC  
TCATATGCTGCTGCAGAATGAGCTCTTCTGAACGTGGGCACCAAGGACTTCCCAGACGGAGAGCTTCGGGGGCA  
ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCCCTGCCCGTGGCCCTAGCAGGAGCCCTGGTGTACC  
CCCTGTGAAGAGCCAAGCAGCAGGGCACGCTGGCTTTCCTTGGATACCCACTGTACCTGCACTATGAAGTGCT  
GCTGGCTGGGCTTGGTGGCTCAGAACAGGCAGTGTCACTGCCACCTCCTTGGGCTCCTGGAACGCCAGGGCC  
TCGGCGGCTGCTGAAGGATTCTATGGCTCAGAGGCCAGGGTGTGGTGAAGGACCTGGAGCCGGAACCTGCTGCG  
GCACCTGGCAAAAGGCATGGCTTCCCTGATGATCACCAACAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCT  
CTCCTCCCAGGTGCACATAGCCAACCAATGTGAGGTGGCGGACTGCGCTGGAGGCGGCCGGGGCCGAGGGGGT  
GCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCCCGCCTGTGGTGCCTGGTCTCCCGGCCCTAGCGCCCCG  
CAAACCTGGTGGTCTTGGGCGGCCCGAGACCCCAACACATGCTTCTTCGAGGGGCGAGCAGCGCCCCACGGGGC  
TCGCTGGGCGCCCAACTAGCACCCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGTCTGTGACCCGGT  
GGTGTGCCACCCGCCCAGCTGCCACACCCGGTGAGGCTCCCGACCAGTGTGCCCTGTTTGGCCTGGCTGCTA  
TTTTGATGGTGACCGGAGCTGGCGGGCAGCGGGTACCGGTGGCACCCCGTTGTGCCCCCTTTGGCTTAATTAA  
GTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCAGTGTCCCCGGCTGGC  
CTGTGCCAGCCTGTGCGTGTCAACCCACCGACTGCTGCAAACAGTGTCCAGGTGAGGCCACCCCCAGCTGGG  
GGACCCCATGCAGCTGATGGGCCCCGGGCTGCCGTTTGTCTGGGAGTGGTTCAGAGAGTCAGAGCTGGCA  
CCCCCTCAGTGCCCCCGTTTGGAGAGATGAGCTGTATCACTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGT  
GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCCAGGGAGGTCCCTGAAGAAGCTGAAGGTCA  
CTGTGTCCCAGTGCTCTGGGGGACACTCAGTGTCTGCTCTGTCTGTACCAGGCAGGGGTGCCTCACTGTGAGC  
GGGATGACTGTTCACTGCCACTGTCTGTGGCTCGGGGAAGGAGAGTGCATGCTGTTCCCGCTGCACGGCCACC  
GGCGGCGTAAGTGAGGGAGTCCAGGGTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCTGTGGGACTCCTGATCAG  
GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCCTCACAAGTGCCATTCCAATCCACCCCTCACAGCA  
ACCTGGTGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTCTGAAGCTCAGAGAAATTAAGCAACGAG  
ATGAAGGTACCCAGCTGTGTGCACTGACCTGTTTAGAAAATACTGGCCTTTCTGGGACCAAGGCAGGGATGCTT  
TGCCCTGCCCTCTATGCCCTCTCTGTGCTCTCCACTCCCTCTCCCTCCTCCAACATTCCCTCCCTTCTGTCTCC  
AGCAGCCCCAGAGACCAGAACTGATCCAGAGCTGGAGAAAGAAGCCGAAGGCTCTTAGGGAGCAGCCAGAGGGCC  
AAGTGACCAAGAGAGTAGGGGCTGAGCTGGGAAGGGGTGGCATCGAGGACCTTCTGCATTCTCCTGTGGGAAG  
CCCAGTGCCCTTTGCTCCTCTGTGCTGCTTACTCCACCCCCACTACCTCTGGGAACCAAGCTCCACAAGGGG  
GAGAGGCAGCTGGGCCAGACCGAGGTACAGCCACTCCAAGTCTGCCCTGCCACCCCTCGGCCTCTGTCTGGAA  
GCCCCACCCCTTCTTCTCTGTACATAATGTCACTGGCTTGTGGGATTTTTAATTTATCTTCACTCAGCACCAG  
GGCCCCGGACACTCCACTCCTGTGCCCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTTGTATTTATTAAAC  
ATTTCTTTTTTCACTCTTTGGGCATGAGGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCTGGTGGAGAAGGG  
GCNAGAGTAGGAGGTGAGAGAGAGGACTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG  
CGTGGCNCNTTGGCTGGCATNCCTGGGTTCCGCGAGAGGGGCTGGGGATGGTTCTTGAGATGGTCTAGAGACTCAAG  
AATTTAGGGAAGTAGAAGCAGGATTTTGACTCAAGTTTAGTTTCCACATCGCTGGCCTGTTGCTGACTTCATG  
TTTGAAGTTGCTCCAGAGAGAGAATCAAAGGTGTACCAGCCCCCTCTCTCCCTCCTTCCCTTCCCTTTCT  
TTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC

093555 073004

## FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA  
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC  
TCCACCATGAACGTGCGGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG  
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC  
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA  
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT  
GGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTGCGA  
GGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC  
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA  
GAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC  
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG  
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTAC  
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG  
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC  
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT  
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAACTGCCCCCT  
CCTTTCTTTCTTCTTTCTTCCCTCCCTCCCTTTCTTCCCCTTTTCTTCCCTTCTTCCCTTCC  
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG  
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTTCATGGACCT  
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTGAGGAGAGACTTATGGAGCC  
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

## FIGURE 228

ATGCCTACTACCTTCCAATAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC  
TTCTATGCCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA  
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC  
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG  
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT  
TGCTGACAACGGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG  
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT  
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC  
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGGC  
ACTTCGGCTGCCCTGTTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC  
TGGATCAGGGGAGAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGT  
TTGCTCTTGGGTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA  
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCCCT  
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG  
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT  
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG  
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC  
TGGCTCACCTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCCAGTG  
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC  
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT  
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA  
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT  
TGGCCCTTATAGGACC

993450

## FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA  
 GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAAACAACACCCAACAACCTGGGGTGG  
 GGGGAAGAAAGAAAGAAAAGAAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAATCCTGTGGCGCGCCGCCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG  
 TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG  
 CTGTTGGTGCAGGGTGCTTGTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG  
 CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCTGGGCGGCCGTGG  
 ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT  
 TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTGCGGGAGGTGATAAGTGGTCAGT  
 GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG  
 TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTTCAGACTCAACATACACCCAGAACA  
 ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC  
 CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGGCCACTGGGAAACCAGAGCCTTCCA  
 TTTCTTGGCGACACATCTCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT  
 TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCAGGAAAATGCTGTGTCAAT  
 CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTTCAGGAAATTAAT  
 CTGGCACCGTGACCCCCGGACGCGAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT  
 CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT  
 TCAAAATTTTAGCACAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA  
 ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT  
 CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT  
 TGTGTTGACACTGTCCTCTTTTACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA  
 TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT  
 GGTACAGTTTGTAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC  
 TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTTGATTAGCTACATTA  
 CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG  
 AGGATATTAATTGTGATTTTCATGTTTGTAAATCTACAACCTTTTCAAAGCATTTCAGTCATGGT  
 CTGCTAGGTTGCAGGCTGTAGTTTACAAAACGAATATTGCAGTGAATATGTGATTCTTTAA  
 GGCTGCAATACAAGCATTTCAGTTCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG  
 CATTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA  
 TAACACATATCTAGATTTTTTCTGCTTGCATGATATTTCAGGTTTCAGGAATGAGCCTTGTAAT  
 ATAACCTGGCTGTGCAGCTCTGCTTCTCTTTCTGTAAGTTCAGCATGGGTGTGCCTTCATAC  
 AATAATATTTTTCTCTTTGTCTCCAATAATAAAATGTTTTGCTAAATCTTACAATTTGA  
 AAGTAAAAATAAACCAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC  
 TATTGGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT  
 AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTT  
 TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAATGACAGCACA  
 GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTTCACTAGTCCAAGCCAAAA  
 TTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA  
 AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA  
 TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC  
 AGGAGGAAATTACAGAGACAATTATGACAACTGAAATGAGACATGCACATAATATAGATACA  
 CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAAGGTTTTTG  
 AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA  
 AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA  
 AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

594555 0300  
 100000 000000

## **FIGURE 230**

MMLLVQGACCSNQWLA AVL LLSLCCLLP SCLPAGQSVDFPWA AVDNMMVRK GDTAVLR CYLED  
GASKGAWLNRSSII FAGGDKWSV DPRVSI STL NKR DYSLQIQ NV DVTDDGPYTCSVQTQHTP  
RTMQVHLTVQVPPKIYDISNDMTVNEGTVNLTLCLATGKPEPSISWRHISPSAKPFENGQYL  
DIYGITRDQAGEYECSAENAVSFPDVRKV KVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV  
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL  
NPPSTAQYGITGSADVL FSCWYLVLTLS SFTSIFYLKNAILQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-31

#### **Transmembrane domain:**

amino acids 326-345

#### **N-glycosylation sites.**

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

#### **Casein kinase II phosphorylation site.**

amino acids 147-151, 208-212, 224-228

#### **Tyrosine kinase phosphorylation site.**

amino acids 178-186

#### **N-myristoylation sites.**

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,  
302-308, 319-325

#### **Myelin P0 protein:**

amino acids 92-121

## FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTGCTGG  
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA  
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTACCAAATTGCAATGGAGCCTTTTGAAA  
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTCTCCCTAGCTGTGGTGGTCATC  
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC  
GCGGCTCCGGGTCTGGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT  
CCTTCTCCTTGCTGCAGTCAGCACACCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG  
CAAGTCCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT  
AGACAACCTTCACTCAGAACCCAGGGATGTTTCAAGATCAAAGGTGAACAAGGCGCCCCAGGTC  
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGGCCCTGGCCCGCCGGGACCACCTGCT  
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGGACCCCAAGG  
CCCACCGGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG  
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT  
GGCCCCAAAAGGGGAACTGGAACCTAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA  
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCCTCCTGGAGCCAGGGGAGTA  
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTTTTCTTGAGCTAAAGGAGAT  
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGGCCTCCTGGTGCAGTGGGACACCCAGGTGC  
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA  
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA  
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG  
GCTGGCAGGTCCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG  
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC  
GTCAGGATTGTTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG  
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT  
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT  
GTTTCAAGTGTGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA  
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCGGAACCCCTTTCA  
CTTCTCTGCTCCCGAGGTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT  
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAAATAAGCTCAACATCATTGA

094655-0304

## **FIGURE 232**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAPHQIAMEPF EINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL  
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR  
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD  
GATGPSGPGQPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE  
KGDGLGLPGSKGDRGMKGDAGVMGPPGAQGSKGDGFRPGPPGLAGFFGAKGDQGPGLQGVPG  
PPGAVGHPPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDGTLQGGQGRKGESGVPGPA  
GVKGEQGSPLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR  
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWL DNVQCRGTESTLW  
SCTKNSWGHHDCSHEEDAGVECSV

### **Transmembrane domain:**

amino acids 47-66 (type II)

### **N-glycosylation sites.**

amino acids 43-47, 83-87, 136-140

### **Tyrosine kinase phosphorylation site.**

amino acids 432-440

### **N-myristoylation sites.**

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,  
400-406, 441-447, 475-481, 490-496, 515-521

### **Amidation site.**

amino acids 360-364

### **Leucine zipper pattern.**

amino acids 56-78

### **Speract receptor repeat**

amino acids 422-471, 488-519

### **Clq domain proteins.**

amino acids 151-184, 301-334, 316-349

[illegible]

1. The first part of the document is a title page. It contains the title "The Role of the State in the Development of the Economy" and the author's name "John Doe".

2. The second part of the document is an abstract. It provides a brief summary of the main findings and conclusions of the study.

3. The third part of the document is the introduction. It discusses the importance of the state in the development of the economy and the research objectives of the study.

4. The fourth part of the document is the literature review. It examines the existing research on the role of the state in the development of the economy.

5. The fifth part of the document is the methodology. It describes the research methods used in the study, including data collection and analysis techniques.

6. The sixth part of the document is the results and discussion. It presents the findings of the study and discusses their implications for the development of the economy.

7. The seventh part of the document is the conclusion. It summarizes the main findings and conclusions of the study.

8. The eighth part of the document is the references. It lists the sources of information used in the study.

9. The ninth part of the document is the appendix. It contains additional information related to the study, such as data tables and figures.

10. The tenth part of the document is the index. It provides a list of keywords and their corresponding page numbers.

**REPORT OF THE**

&lt;subunit 1 of 1, 98 aa, 1 stop

MKLMVLVFTIGLTLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG

Important features:

amino acids 1-20

amino acids 72-76

amino acids 63-71

## FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCAGCCTGCAGGG  
 CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC  
 GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG  
 TGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCGCGCTGGCTGTG  
 CGCTGGGGCGCTGGTGTGCGGGTGGCTTCTTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA  
 TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT  
 GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC  
 AGGAACAGAACAAAACCTTTCAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTTGGCC  
 TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGTTGTCTACCCAAATAAGACTCATCCC  
 AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC  
 ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTCTCTCCTC  
 AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAA  
 TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT  
 TTTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT  
 CCGACCCTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT  
 GGAGGTGGTGTCCAGCGTGGAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC  
 AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA  
 GTATTCTGTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC  
 TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG  
 CTTTACTGGAAACTTTTCTACACAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA  
 CGAGAATTTACAATTGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT  
 CTGGGAGGTCACCGGACTCATGGGTGTTTGGTGGTATTGACCCCTCAGAGTGAGCAGCTGT  
 TGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGAGACCTAGAAGAA  
 CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCA  
 GAGGAGAATTCAAGACTCCTTCAAGAGCGTGCGGTGGCTTATATTAATGCTGACTCATCTAT  
 AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC  
 TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG  
 ACTAAAAAAGTCCTTCCCCAGAGTTCAGTGGCATGCCAGGATAAGCAAATTGGGATCTGG  
 AAATGATTTTGGAGGTGTTCTTCCAACGACTTGAATTTGCTTCAGGCAGAGCACGGTATACTA  
 TAAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT  
 GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG  
 AGGAGGGATGGTGTGTTGAGCTAGCCAATTCCATAGTGCTCCCTTTTGATTGTGAGATTATG  
 CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA  
 ATGAAGACATACAGTGTATCATTGATTCACTTTTTCTGCAGTAAAGAATTTTACAGAAAT  
 TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA  
 TGATGAATGATCAACTCATGTTTCTGGAAAAGCATTTATTGATCCATTAGGGTTACCAGAC  
 AGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC  
 ATCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGACCCCTTCCAAGGCCT  
 GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT  
 TTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA  
 CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAATTTGGTATATTTGAAATAAAGT  
 TGAATATTATATATAA

## **FIGURE 236**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL  
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH  
PNYISIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF  
KLERDMKINCSGKIVARYGKVFRGNKVKNALAGAKGVILYSDPADYFAPGVKSYPDGWNL  
PGGGVQQRGNILNLNAGADPLTPGYPANAYRRGIAEAVGLPSIPVHPIGYYDAQKLEKMG  
GSAPPDSSWRGSLKVPYNVGPFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV  
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW  
AEENSRLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES  
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET  
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ  
EMKTYSVSFDSLFSVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP  
DRPFYRHVIYAPSSHKNKYAGESFPGIYDALFDIESKVDPSKAWGEVQRQIYVAAFTVQAAAE  
TLSEVA

### **Signal sequence:**

amino acids 1-40

### **N-glycosylation sites.**

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,  
459-463, 476-480, 638-642

### **Tyrosine kinase phosphorylation sites.**

amino acids 363-372, 605-613, 606-613, 617-626

### **N-myristoylation sites.**

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,  
360-366, 427-433, 529-535, 707-713